Universidad Autónoma de Madrid

Probing the Dark Universe with Gravitational Waves

Jose María Ezquiaga Bravo

Doctoral Thesis in Theoretical Physics at UAM, Spain 2019

Universidad Autónoma de Madrid

Probing the Dark Universe with Gravitational Waves

Jose María Ezquiaga Bravo

Memoria de tesis doctoral presentada ante el Departamento de Física Teórica de la Universidad Autónoma de Madrid para la obtención del título de Doctor en Física Teórica.

Tesis Doctoral dirigida por

Juan García-Bellido Capdevila

Catedrático en Física Teórica de la Universidad Autónoma de Madrid.

Madrid, junio de 2019

para Ana

Abstract

Gravitational wave (GW) astronomy opens new opportunities to explore the universe and its fundamental laws. This thesis focuses on probing the pillars of the standard cosmological model with GWs, specially its most puzzling components: dark energy (DE) and dark matter (DM). We propose and apply new tests of DE and General Relativity (GR) with the propagation of GWs. We also investigate the formation of black-holes (BHs) in the early universe, which has strong implications on their contribution to the DM and on their GW signatures.

Just as electromagnetic radiation can scan materials, GWs can probe the medium in which they propagate. DE models beyond Einstein's gravity generically modify the propagation of GWs. We identify the speed of GWs as a key test of gravity and find the conditions for an anomalous speed to arise. We emphasis that a non-luminal speed can appear in cosmological models aiming at DE such as Galileons, but also in environments with a spatial profile induced by screening or scalar hair. After the multi-messenger event GW170817, we determine the consequences of the tight constraint on the speed of GWs for different classes of gravity theories and DE models, setting the dead ends and the road ahead. Standard sirens like GW170817 constrain as well the GW luminosity distance. We derive this observable in general theories of gravity and discuss its detectability with the future space-based detector LISA. Particularly distinguishable oscillatory patters are produced by GW oscillations, a phenomenon that we study in detail. Other probes of GW oscillations are modified wave-forms, induced anomalous speeds and polarization dependent signals.

Primordial BHs (PBHs) could be a unique relic to unveil the physics of the early universe. We study the production of PBHs in single field model of inflation with a quasi-inflection point, showing the growth of perturbations beyond slow-roll (SR) at sub- and super-horizon scales. We propose a particle physics motivated model, critical Higgs inflation, achieving a copious PBH production with several GW signatures. However, when curvature fluctuations are enhanced, quantum diffusion dominates the classical inflationary dynamics. We develop a formalism based on stochastic inflation beyond SR to account for this effect. We encounter that the classical prediction is importantly modified, with relevant non-Gaussian contributions. To quantify better the quantum correction, we devise a method to compute directly the tail of the curvature perturbation distributions. As a first step, we apply it to SR inflation. We conclude that the abundance of PBHs is many orders of magnitude larger than the Gaussian prediction, discussing its implications for inflationary model building as well as for the GW observables.

Altogether, GW astronomy stands as a powerful channel to advance forward in the quest for understanding the dark universe. We discuss the future prospects of this line of research, highlighting the theoretical challenges and observational opportunities that next generation GW detectors will provide.

Resumen

La astronomía de ondas gravitacionales (OG) abre nuevas oportunidades para explorar el universo y sus leyes fundamentales. Esta tesis se centra en investigar los pilares del modelo cosmológico estándar con OG, especialmente sus componentes más desconocidas: la energía oscura (EO) y la materia oscura (MO). En esta tesis proponemos y aplicamos nuevas pruebas de la EO y la Relatividad General (RG) con la propagación de las OG. También investigamos la formación de agujeros negros (AN) en el universo primitivo, así como sus implicaciones para la MO y OG.

Al igual que las ondas electromagnéticas pueden escanear materiales, las OG pueden sondear el medio en el que se propagan. De manera genérica, los modelos de EO más allá de la gravedad de Einstein modifican la propagación de las OG. Nosotros identificamos la velocidad de las OG como una prueba clave de la gravedad y encontramos las condiciones para que surja una velocidad anómala. Es importante hacer hincapié en que una velocidad distinta a la de la luz puede aparecer en modelos cosmológicos como Galileons, pero también en entornos con un perfil espacial inducido por el screening o el scalar hair. Tras del evento multi-mensajero GW170817, grandes familias de teorías de la gravedad y modelos de EO quedaron descartados. Sirenas estándar como GW170817 sirven también para medir la distancia luminosidad de las OG. Aquí calculamos su posible modificación respecto de RG y discutimos su posible observación con el futuro detector LISA. Las oscilaciones de OG, un fenómeno que estudiamos en detalle, también pueden dejar señales distinguibles en la distancia luminosidad. Otras pruebas de las oscilaciones de OG son modificaciones de la wave-form, velocidades anómalas y dependencia en la polarización.

Los AN primordiales (ANP) podrían ser una reliquia única para revelar la física del universo primitivo. Estudiamos la producción de ANP en un modelo de campo único de inflación con un punto de inflexión, produciendo un crecimiento de las perturbaciones más allá de slow-roll (SR) en escalas de sub- y súper-horizonte. Un modelo motivado por la física de partículas que nosotros proponemos es critical Higgs inflation, que logra una producción abundante de ANP con varias señales de OG. Sin embargo, cuando se incrementan las fluctuaciones de curvatura, la difusión cuántica domina la dinámica inflacionaria clásica. Por tanto, desarrollamos un formalismo basado en la inflación estocástica más allá de SR para dar cuenta de este efecto. Encontramos que la predicción clásica se modifica de manera importante, con contribuciones relevantes no Gaussianas. Para cuantificar mejor la corrección cuántica, ideamos un método para calcular directamente la cola de las distribuciones de perturbación de curvatura. Como primer paso, lo aplicamos a la inflación de SR. Concluimos que la abundancia de PBH es muchos órdenes de magnitud más grande que la predicción Gaussiana, discutiendo sus implicaciones para modelos inflacionarios así como para los observables de OG.

En total, la astronomía de OG se erige como un canal poderoso para avanzar en la comprensión del lado oscuro del universo. Para concluir, presentamos nuestras perspectivas futuras de este campo, destacando los desafíos teóricos y las oportunidades observacionales que proporcionarán los detectores de OG de próxima generación.

Declaration

This thesis is the result of four years of work. During this time, different themes and topics have been explored. Some of the results presented here have already been published or appeared as preprints. As of June 17, 2019, the complete list of publication in chronological order is[1]

[1] *Towards the most general scalar-tensor theories of gravity: a unified approach in the language of differential forms*
JM. Ezquiaga, J. García-Bellido and M. Zumalacárregui
Phys.Rev.D 94, 024005 (2016) (arXiv:1603.01269)

[2] *Speed of Gravitational Waves and the Fate of Scalar-Tensor Gravity*
D. Bettoni, **JM. Ezquiaga**, K. Hinterbichler and M. Zumalacárregui
Phys.Rev.D 95, 084029 (2017) (arXiv:1608.01982)

[3] *Field redefinitions in theories beyond Einstein gravity using the language of differential forms*
JM. Ezquiaga, J. García-Bellido and M. Zumalacárregui
Phys.Rev.D 95, 084039 (2017) (arXiv:1701.05476)

[4] *Primordial Black Hole production in Critical Higgs Inflation*
JM. Ezquiaga, J. García-Bellido and E. Ruiz Morales
Phys.Lett.B 776. 345-349 (2018) (arXiv:1705.04861)

[5] *Dark Energy after GW170817: dead ends and the road ahead*
JM. Ezquiaga and M. Zumalacárregui
Phys.Rev.Lett. 119.251304 (2017) (arXiv:1710.05901)

[6] *Quantum diffusion beyond slow-roll: implications for primordial black-hole production*
JM. Ezquiaga and J. García-Bellido
JCAP 2018 (2018) no.08, 018 (arXiv:1805.06731)

[7] *Dark Energy in light of Multi-Messenger Gravitational-Wave astronomy*
JM. Ezquiaga and M. Zumalacárregui
Frontiers in Astronomy and Space Sciences (arXiv:1807.09241)

[8] *Testing modified gravity at cosmological distances with LISA standard sirens*
E. Belgacem *et al.* (including **JM. Ezquiaga**)
Submitted to JCAP (arXiv:1906.01593)

[1] For an updated list of publications visit http://inspirehep.net/author/profile/J.M.Ezquiaga.1.

The thesis is also based on the following ongoing projects

[9] *Speed of Gravitational waves over scalar hair and cosmologically trivial solutions*
JM. Ezquiaga and M. Zumalacárregui

[10] *Implications of gravitational waves oscillations for Dark Energy*
J. Beltrán Jiménez, **JM. Ezquiaga** and L. Heisenberg

[11] *Exploring the tail of curvature perturbation distributions to properly assess PBH abundance*
JM. Ezquiaga, J. García-Bellido and V. Vennin

Each of these works have contributed differently to the content of the thesis. A brief summary of the relation with each chapter would be:

- Chapter 1 and 2 are mainly based on the review of gravitational wave tests of dark energy [7].
- Chapter 3 is based on the derivation of the conditions for an anomalous speed [2], the implications of GW170817 for dark energy [5] and the study of the spatial delay in non-cosmological backgrounds [9].
- Chapter 4 is based in part of the derivation of the GW luminosity distance in theories with an anomalous GW speed performed in [8].
- Chapter 5 is based mostly on the study of GW oscillations in [10] except the last section which is based on the study of the detectability go bigravity GW oscillation developed for the LISA Cosmology Working Group [8], which makes use of the LISA massive black-hole binaries catalogue.
- Chapter 6 is based in part on the analysis of the enhancement of the power spectrum in inflationary potential with an inflection point [6].
- Chapter 7 is based on the study of the primordial black-hole production in critical Higgs inflation [4].
- Chapter 8 is based mostly on the investigation of the implications of quantum diffusion for the production of PBH [6].
- Chapter 9 is based on the study of the tail of the curvature perturbation distributions and their implications to the PBH abundance[11].
- Chapter 10 and 11 are a summary of the results and future prospects of all the investigations during the thesis.
- Appendices A, B and C are based on [1, 3], [2] and [6] respectively.

My research has been supported by the Spanish FPU Grant No. FPU14/01618; the Research Projects FPA2013-47986-03-3P and FPA2015-68048-03-3P (MINECO-FEDER); and the Centro de Excelencia Severo Ochoa Programs SEV-2012-0249 and SEV-2016-0597.

Acknowledgments

The path towards this thesis has been a beautiful journey. I can feel only lucky and thankful for all the opportunities and support I received. Countless actions and people have made this endeavor possible, both at the academic and personal level. So let me begin with a global thank you to all of you who contributed in this adventure.

Undoubtedly, the people who have had more impact on my scientific growth are *Juan García-Bellido*, my thesis advisor, and *Miguel Zumalacárregui*, my closest collaborator. Everything would have been different without your advice, knowledge and enthusiasm. We enjoy doing science together, and this is the best gift.

I would like to thank also to my *collaborators*, whose ideas and expertiese have helped to shape my research perspectives and skills. Science is a collaborative activity and you are all part of the results of this thesis. In the same manner, I am very grateful to my *academic siblings* for their guidance at different stages.

Of course a doctorate is as well a personal and emotional challenge that is much easier to pass with the help of others. *Ana*, you can take most of the credit here. *Family*, *friends* your support was indispensable. Gracias.

Contents

Part I

Introduction

CHAPTER

1

Cosmology in a nutshell

The Standard Model of Cosmology (or ΛCDM) stands as a robust description of our universe. It is based on the theory of General Relativity (GR), which dictates the long-range gravitational interactions, together with the Cosmological Principle, which describes the geometry as homogeneous and isotropic on large scales. Standard matter (baryons, photons, neutrinos...) represents only a small fraction of the energy budget of the universe. The main ingredient is dark energy (DE), an unknown substance causing the late time acceleration. The other major component is dark matter (DM), an undetected constituent that seeds cosmic structures. The last piece of the Standard Model (SM) of Cosmology are the initial conditions, which are thought to be set by an early period of quasi-exponential expansion known as inflation. Despite the observational success of this model [12], it remains as a puzzle the fundamental origin of each piece, which could be associated to new physics as we schematically summarize in Fig. 1.1. The aim of this thesis is to explore the dark sectors of the universe using gravitational waves.

In the SM of Cosmology, the current accelerated expansion is explained by a constant energy density acting as a perfect fluid with negative pressure. Such a cosmological constant (CC) term is perfectly consistent with present observations but notoriously disagrees with theoretical expectations for the vacuum energy [13, 14]. If this energy density is let to evolve in time, one naturally arrives to a dynamical description of DE sourced by a cosmological scalar field [15]. If this field is now allowed to interact (non-minimally) with gravity, the possibilities to describe the cosmic expansion escalate [16]. Alternatives to ΛCDM offer the possibility to alleviate some of its tensions. For instance, DE models with an effective equation of state more negative than the cosmological constant could ease the tension between the local measurement of the Hubble constant and the inferred value from the cosmic microwave background (CMB). Exploring the largest scales with galaxy surveys like Euclid or LSST will help us understanding the expansion history of the universe and will provide new insights about gravity.

Gravity can be tested at different scales and regimes. Classical tests of gravity range from laboratory experiments to Solar System distances, and cover gravity in its weak field regime [17]. Astrophysical observations provide new avenues to improve these tests [18]. Pulsars in particular can be especially constraining, for instance with the recent observations of a triple stellar system [19]. Tests in a much stronger regime have been performed tracking stellar orbits around the galactic center [20]. Altogether, these observations severely constrain modifications of GR. Theories beyond Einstein's theory

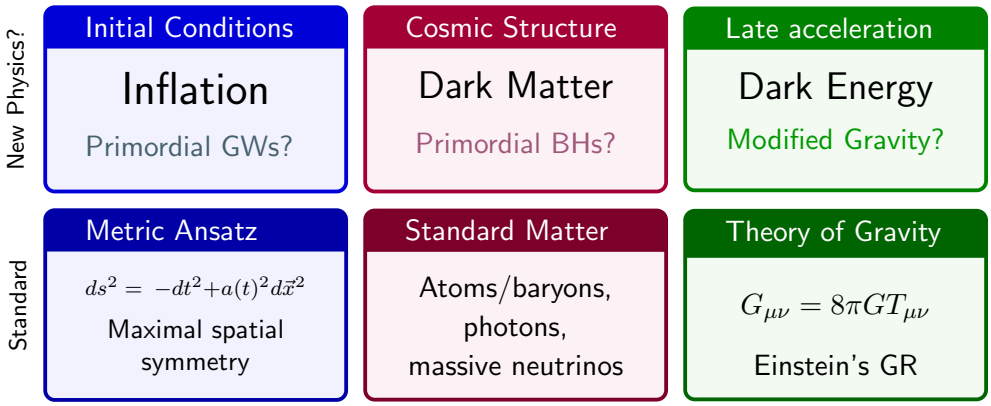

Figure 1.1. Schema of the ingredients of the Standard Model of Cosmology and their possible connection with new physics. In this thesis we will study how to probe the dark side of the universe, dark energy and dark matter, with gravitational waves. In part II, we will investigate modify gravity models aiming at explaining dark energy. In part III we will analyze the formation of PBHs in the early universe and their implications for the cosmic structures.

should thus resemble GR at small scales, e.g. hiding fifth forces with screening mechanisms [21, 22]. At large scales, however, present constraints are considerably weaker. Combining different probes could be crucial to set an observational program to test gravity from cosmology [23].

Gravitational wave (GW) astronomy offers the possibility to test gravity both in the strong regime and at large scales. So far, in the first two runs of the advanced detectors, there have been eleven individual detections, ten binary black-holes (BBH) and one binary neutron star (BNS) [24].[1] No GW background [25], periodic source [26] or long-duration transient [27] have been detected.

GWs could be critical in resolving the open problems of the SM of Cosmology. For instance, (non) observations of cosmological backgrounds of primordial GWs test inflation. BBHs events teach us about the population of BHs, which constrains their possible contribution to DM and their possible primordial origin [28]. Moreover, if DM is described by ultra-light bosons or axions, it could resonate with pulsars [29] or form clouds around BHs observable with GWs [30]. Finally, DE can be probed with multi-messenger events with an associated counterpart such as GW170817 [31, 32] that become standard sirens [33] or with GW events and a statistical analysis [34] being dark sirens. In this work we will focus on this last case, exploring the possibilities of multi-messenger GW astronomy to probe the nature of DE and the fundamental properties of gravity (see a schematic timeline of present and future facilities in Fig. 2.7).

[1] During the completion of this thesis the third observing run was ongoing. Several candidates were publically announced. For an updated list of all the candidate events see https://gracedb.ligo.org/latest/.

1.1 Dark energy and modified gravity

The enigma of the late time acceleration is intimately related to our knowledge of gravity at large scales. In this section we are going to survey possible extensions of Einstein's theory of gravity relevant for DE. We will begin with a theoretical approach to generalize GR in section 1.1.1, to end with phenomenological descriptions of gravity at cosmological scales in section 1.1.2.

1.1.1 Theories of Gravity

Gravity can be modified in different ways. A common path is to generalize the Einstein-Hilbert action

$$S_{GR} = \int d^4x\sqrt{-g}\frac{R[g_{\mu\nu}]}{16\pi G} + S_m[g_{\mu\nu}, \cdots], \tag{1.1.1}$$

where G is Newton's constant and S_m denotes the action of matter, universally and minimally coupled to the metric $g_{\mu\nu}$. Variation of the action (1.1.1) with respect to the metric leads to Einstein's field equations

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G T_{\mu\nu}, \tag{1.1.2}$$

where $R_{\mu\nu}$ is the Ricci tensor, $R \equiv g^{\mu\nu}R_{\mu\nu}$ the Ricci scalar and $T_{\mu\nu} = \frac{-2}{\sqrt{-g}}\frac{\delta S_m}{\delta g^{\mu\nu}}$ is the matter energy-momentum tensor. Einstein's equations can be used to obtain solutions for the space-time ($g_{\mu\nu}$) given the matter content ($T_{\mu\nu}$) in any physical situation, including cosmological solutions relevant to study dark energy.

The structure of gravitational theories is severely restricted and several results can be used to prove the uniqueness of General Relativity under quite broad assumptions. Weinberg's theorems restrict possible infrared (low energy) interactions of massless, Lorentz invariant particles, which for spin-2 lead unavoidably to the equivalence principle [35] and the derivation of Einstein's equations [36].[2] At the classical level, the results of Lovelock imply that the Einstein-Hilbert action is unique in 4D [38, 39].

According to the above results, alternative theories of gravity can be classified into those that

- Break the fundamental assumptions.
- Include additional fields.
- Make the graviton massive.

Note that those descriptions are not exclusive, and many theories fall within several categories. For instance: bimetric gravity has an additional field (tensor) and contains a massive graviton, Einstein-Aether is both Lorentz-violating and includes a vector field, TeVeS has a scalar in addition to a vector, and many extra-dimensional models can be described in terms of additional fields in certain limits. Also, when referring to massive gravitons, we will be considering only classical spin-2 fields.

[2] In addition to GR, there is another theory for massless, spin-2 fields in 4D, Unimodular Gravity, which is invariant under diffeomorphisms preserving the 4D volume element [37].

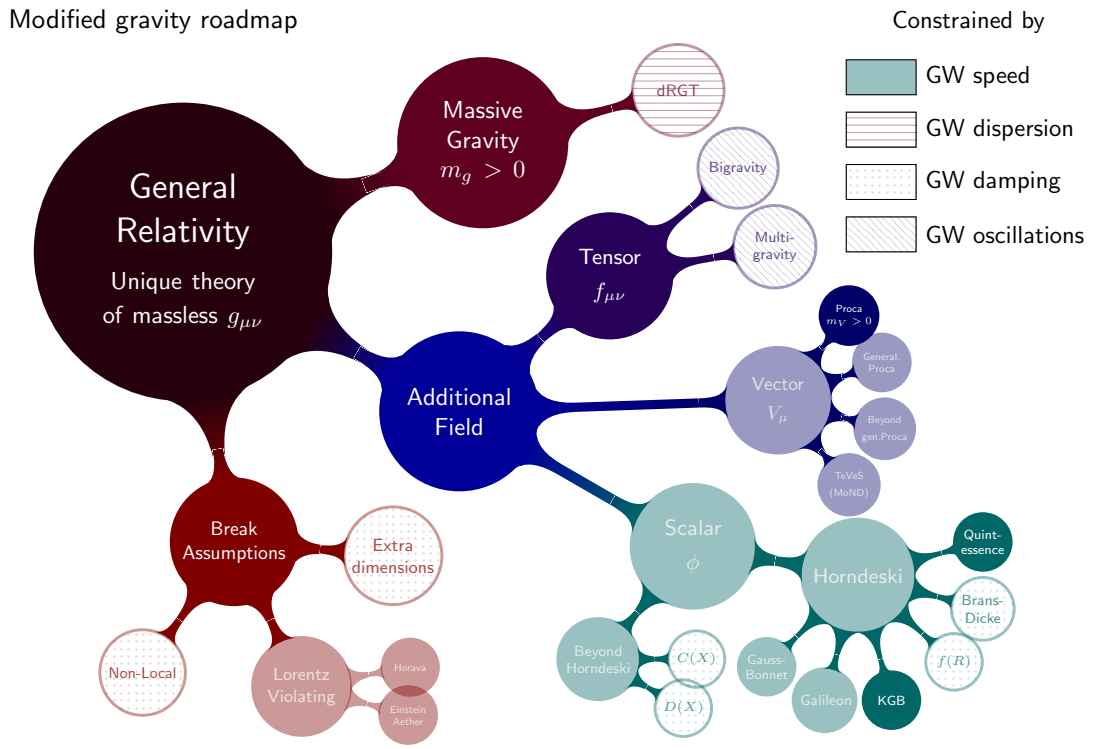

Figure 1.2. Modified gravity roadmap summarizing the possible extensions of GR described in Sec. 1.1. The main gravitational wave (GW) test of each theory is highlighted. For details in the different tests see the discussion in chapter 3 (GW speed and dispersion), 4 (GW damping) and 5 (GW oscillations). Theories constrained by the GW speed and GW oscillations can also be tested with GW damping and GW dispersion respectively. Note in addition that many theories fall under different categories of this classification (see text in Sec. 1.1.1).

Breaking fundamental assumptions

The theorems that fix the structure of General Relativity assume a four dimensional pseudo-Riemannian manifold and local interactions satisfying Lorentz invariance. Any departure from these principles offers a way to construct modified theories of gravity.[3]

Extra dimensions: Additional spatial dimensions allow the inclusion of new operators constructed only from the metric tensor. The canonical examples are Lovelock invariants [38], such as the Gauss-Bonnet term (a topological term in 4 dimensions which does not contribute to the equations of motion). The lack of observations of extra dimensions requires some mechanism to hide them. One example is compactification, when extra dimensions are sufficiently small that they are not accessible to experimental tests [40]. A radically opposite view consist on Braneworld constructions, in which the standard model fields live in a 3+1 dimensional brane, embedded in the higher dimensional space [41].

[3] A class of GR extensions include additional geometric elements like torsion or non-metricity. These elements can be viewed as either breaking the fundamental assumptions or including additional fields.

The Dvali-Gabadadze-Porrati (DGP) model [42, 43] is one such construction in which self-accelerating solutions[4] can be obtained. However, this branch of solutions is plagued by a ghost instability. The 4D effective theory can avoid this problem and was the origin of Galileon gravity [44].

Lorentz Invariance Violation: Gravity can be extended by breaking Lorentz invariance. In many of these alternatives a preferred time direction emerges spontaneously breaking Lorentz symmetry (see [45] for a review). Hořava gravity [46] implements Lorentz violation through a preferred foliation of space-time, with the attractive property that Lorentz symmetry can be recovered at low energies Another class of Lorentz-violating theories is Einstein-Aether, in which a vector field with constant norm introduces a preferred direction [47]. The special case of Einstein-Aether theories in which the vector field is the gradient of a scalar is known as Khronometric [48]. Khronometric theories describe the low-energy limit of some extension of Hořava-gravity, linking the two frameworks [49]. These ideas have been studied as cosmological scenarios [50, 51].

Non-local theories: Non-local theories include inverse powers of the Laplacian operator. These models can involve general functions (e.g. $R \cdot f(\Box^{-1}R)$) [52, 53] or be linear (e.g. $R\frac{m^2}{\Box^2}R$) [54]. The latter class of models lead to phantom dark energy [55, 56] and are compatible with cosmological observations (see [56] for a review). However, their viability on the solar system is disputed due to the time evolution of the effective degrees of freedom and the lack of a screening mechanism [57]. Non-local interactions have been also proposed as a means to improve the ultra-violet behavior of gravity, see e.g. [58]. Non-local models are constructed using the Ricci scalar, since non-local terms involving contractions of the Ricci tensor give rise to cosmological instabilities [59, 60].

Additional fields

Gravity can be extended by the inclusion of additional fields that interact directly with the metric. These theories will vary by the type of field (scalar, vector, tensor) and its interactions with gravity. Theories with additional tensors (bigravity and multigravity) are extensions of massive gravity and will be described in Sec. 1.1.1. We will assume a minimal universal coupling of matter to the metric. For a very complete review of gravity theories containing additional fields, see Ref. [61].

Scalar field A scalar is the simplest field by which gravity can be extended. Scalars do not have a preferred orientation and thus a macroscopic, classical state can exist in the universe without affecting the isotropy of the space-time if it depends only on time. Moreover, a potential term can mimic a cosmological constant very closely in the limit in which the field is varying very slowly (e.g. if the potential is very flat), which is the foundation of the simplest single-field inflation and dark energy models (quintessence). Scalar fields may also arise in effective descriptions of fundamental theories belonging to other categories, such as braneworld constructions [62–64]. These properties had led to

[4] Self-accelerating solutions are those in which there is a late time acceleration without a cosmological constant ($\Lambda = 0$).

a proliferation of scalar-based models to describe accelerating cosmologies, both in the context of inflation and dark energy.

Recent efforts to study scalar-tensor theories have led to a classification based on the highest-order derivatives of the additional field present in the action and the equations of motion, with three generations of theories

1. *Old-school scalar tensor theories:* 1^{st} order derivatives in the action, 2^{nd} order in equations.

2. *Horndeski theories* [65]*:* 2^{nd} order derivatives in the action and 2^{nd} order in equations.

3. *Beyond Horndeski:* 2^{nd} order derivatives in the action and higher order in equations.

The classification is motivated by Ostrogradski's theorem, which states that theories with second and higher (time) derivatives in the action generically introduce unstable degrees of freedom [66,67]. While most physical theories belong to the first class, known loopholes to Ostrogradski's theorem exits, for instance in effective or non-local theories (in which the ghost degrees of freedom are removed) [68] or when the theory is degenerate (that is, the inversion to canonical variables is not possible). The degeneracy condition is automatically satisfied if the equations of motion are second order, but that is not strictly necessary (different conditions appear when there are additional degrees of freedom [69]).Known viable beyond Horndeski theories are known as Degenerate Higher Order Scalar Tensor (DHOST) [70], which have second derivatives in the action (higher derivatives in the equations), but recently toy models with higher derivatives in the action have been proposed [71].

Interestingly, scalar-tensor theories can be reformulated in terms of differential forms in which the second order equations follow naturally from the antisymmetry of this language [1]. This approach can be generalized to gravity theories with additional vector and tensor fields as well [3], introducing field redefinitions to interconnect them. In appendix A we summarize the main aspects of this alternative formulation.

Old-school scalar-tensor theories contain at most first derivatives of the scalar in the action. They can be seen as a generalization of the Jordan-Brans-Dicke theory of gravity [72]

$$S = \int d^4x\sqrt{-g}\frac{M_{\text{Pl}}^2}{2}\left[\omega(\phi)R - K(X,\phi)\right] + S_m\,, \tag{1.1.3}$$

where $X \equiv -\nabla^\nu\phi\nabla_\nu\phi/2$ is the canonical kinetic term of the scalar field. This theory includes GR ($\omega = 1, K = \Lambda$), quintessence ($\omega = 1, K = X - V$) [73, 74], Brans-Dicke models [72] ($\omega = \phi$, $K = \frac{\omega_{\text{BD}}}{\phi}X - V(\phi)$), k-*essence* [75, 76] ($\omega = 1$, $K = K(\phi, X)$). Archetypal modified-gravity models such as $f(R)$ [77–79] are equivalent to instances of these theories [80]. Chameleons [81] and symmetrons [82] also belong to this class of theories. Certain freedom exists in writing the theory due to the possibility of rescaling the metric $g_{\mu\nu} \to \bar{g}_{\mu\nu} = C(\phi)g_{\mu\nu}$ and redefining the scalar field, i.e. the Jordan frame in which the metric is minimally coupled (1.1.3) and the Einstein frame in which ω

is constant but matter is explicitly coupled to the scalar [83]. Current cosmological observations constrain the Brans-Dicke parameter $\omega_{\rm BD} > 692$ (99%) [84].

Horndeski's theory contains the best understood examples of scalar-tensor theories. The Horndeski action encompasses all local, 4D Lorentz invariant actions whose metric and field variation leads to second order equations of motion [65] (Horndeski's theory is also known in the literature as Generalized Galileons [85, 86]). Horndeski's action reads

$$S = \int d^4x\sqrt{-g}\sum_{i=2}^{5}\mathcal{L}_i[\phi, g_{\mu\nu}] + S_m[\chi_i, g_{\mu\nu}], \tag{1.1.4}$$

where we have assumed minimal and universal coupling to matter in S_m. The sum is over the four Lagrangians

$$\mathcal{L}_2 = K(X,\phi), \tag{1.1.5}$$

$$\mathcal{L}_3 = -G_3(X,\phi)\Box\phi, \tag{1.1.6}$$

$$\mathcal{L}_4 = G_4(X,\phi)R + G_{4X}\left\{(\Box\phi)^2 - \phi_{\mu\nu}\phi^{\mu\nu}\right\}, \tag{1.1.7}$$

$$\mathcal{L}_5 = G_5(X,\phi)G_{\mu\nu}\phi^{\mu\nu} - \frac{1}{6}G_{5X}\{(\Box\phi)^3 - 3\phi^{\mu\nu}\phi_{\mu\nu}\Box\phi + 2\phi^{\mu}_{\nu}\phi^{\nu}_{\alpha}\phi^{\alpha}_{\mu}\}, \tag{1.1.8}$$

where K and G_A are functions of ϕ and $X \equiv -\nabla^\nu\phi\nabla_\nu\phi/2$, and the subscripts X and ϕ denote partial derivatives. Horndeski theories include all the generalized Jordan-Brans-Dicke type, plus new additions that involve second derivatives of the scalar at the level of the action. These include kinetic gravity braiding (KGB) $(K(X), G_3(X))$ [87–89], covariant galileons ($K, G_3 \propto X$, $G_4, G_5 \propto X^2$) [44, 90], disformal [91] and Dirac-Born-Infeld gravity ($G_i \propto \sqrt{1 + X/\Lambda_i^4}$) [62, 92], Gauss-Bonnet couplings [1] and models self-tuning the cosmological constant [93, 94]. Just as Brans-Dicke is invariant under rescalings of the metric, Horndeski theories are invariant under field-dependent disformal transformations $g_{\mu\nu} \to \bar{g}_{\mu\nu} = C(\phi)g_{\mu\nu} + D(\phi)\phi_{,\mu}\phi_{,\nu}$, which amount to a redefinition of the Horndeski functions G_i (and the introduction of an explicit coupling to matter) [95].

Theories beyond Horndeski have higher order equations of motion without including additional degrees of freedom. The first examples of these theories [96] were related to GR by a metric redefinition involving derivatives of the scalar field [97],

$$g_{\mu\nu} \to \bar{g}_{\mu\nu} = C(X,\phi)g_{\mu\nu} + D(X,\phi)\phi_{,\mu}\phi_{,\nu}\,, \tag{1.1.9}$$

applied to the gravity sector. The simplest such beyond Horndeski theory emerged from the metric rescaling with derivative dependence $C = \Omega^2(X,\phi), D = 0$, and was dubbed *kinetic conformal gravity* [96]

$$S_C = \int d^4x\frac{\sqrt{-g}}{16\pi G}\left(\Omega^2 R + 6\Omega_{,\alpha}\Omega^{,\alpha}\right) + S_\phi + S_M\,, \tag{1.1.10}$$

where S_ϕ is an additional scalar field Lagrangian. One of the premises in constructing this type of theory was the existence of an inverse for the relation (1.1.9), which can

be studied through the Jacobian of the mapping [96]. If this assumption is broken the resulting theory is *mimetic gravity* [98], a gravitational alternative to dark matter. Interestingly, the conformal relation between kinetic conformal gravity (1.1.10) and GR ensures that this is one of the theories in which the speed of GWs is nontrivially equivalent to the speed of light [5, 99].

The best known beyond Horndeski theory is given by the Gleyzes-Langlois-Piazza-Vernizzi (GLPV) action [100], which consists of Horndeski plus the additional Lagrangian terms:

$$\mathcal{L}_{4b} = B_4(\phi, X)\epsilon^{\mu\nu\rho}_{\sigma}\epsilon^{\mu'\nu'\rho'\sigma}\phi_\mu\phi_{\mu'}\phi_{\nu\nu'}\phi_{\rho\rho'}\,, \tag{1.1.11}$$

$$\mathcal{L}_{5b} = B_5(\phi, X)\epsilon^{\mu\nu\rho\sigma}\epsilon^{\mu'\nu'\rho'\sigma}\phi_\mu\phi_{\mu'}\phi_{\nu\nu'}\phi_{\rho\rho'}\phi_{\sigma\sigma'}\,. \tag{1.1.12}$$

Horndeski and GLPV Lagrangians of the same order, i.e. $\mathcal{L}_4 + \mathcal{L}_{4b}$ (A.1.25+1.1.11) or $\mathcal{L}_5 + \mathcal{L}_{5b}$ (A.1.26+1.1.12), can be mapped to Horndeski via $g_{\mu\nu} \to \hat{g}_{\mu\nu} = C(\phi)g_{\mu\nu} + D(X, \phi)\phi_\mu\phi_\nu$ showing the viability of these combinations [100, 101]. For generic combinations of Horndeski and GLPV, viability arguments were first based on a special gauge (unitary gauge) that assumed that the scalar field derivative ϕ_μ is timelike. Subsequent analyses eventually lead to covariant techniques to study the degeneracy conditions [70] (see [102] for earlier criticism). These techniques later showed that not all Horndeski and GLPV combinations met the degeneracy condition on a covariant level [103].

The study of degeneracy conditions for scalar-tensor theories ultimately led to the *degenerate higher-order scalar-tensor* (DHOST) [70] paradigm classification of theories with the right number of degrees of freedom (also known as Extended Scalar-Tensor or EST) [104]. DHOST theories include cases beyond conformal kinetic gravity (1.1.10) and GLPV theories (1.1.11,1.1.12). DHOST theories are invariant under general disformal transformations (1.1.9), which can in turn be used to classify them [105] (see also [106]). DHOST theories have been fully identified including terms with up to cubic second-field derivatives in the action, e.g. $\sim (\Box\phi)^3$ [107]. Demanding the existence of a Poisson-like equation for the gravitational potential restricts the space of DHOST theories to those that are related to Horndeski via disformal transformations (1.1.9) [108].

Vector field Theories with vector fields have been proposed as modifications to GR and in the context of dark energy. A background vector field does not satisfy the isotropy requirements of the cosmological background, unless it points in the time direction and only depends on time $A_\mu = (A_0(t), 0, 0, 0)$. Isotropy can also happen on average, if a vector with a space-like projection oscillates much faster than the Hubble time [109]. In that case the background is isotropic on average but the perturbations (including gravitational waves) inherit a residual anisotropy [110]. Finally, theories with multiple vectors can satisfy isotropy, for instance, if they are in a triad configuration $A^a_\mu = A(t)\delta^a_\mu$ [111].[5] A large number of vectors can also lead to statistical isotropy (e.g. if the orientations are random) [113]. The kinetic term for a vector field, $F_{\mu\nu}F^{\mu\nu}$, is defined

[5] Technically speaking, multiple vectors can lead to isotropic solutions if they have an internal symmetry that together with the broken space-time symmetries leaves a residual $ISO(3)$ [112]. For the case of the triad, the symmetry group is $SO(3)$.

by the gauge invariant field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and the addition of a mass term $m^2 A_\mu^2$ is known as Proca theory [114].

Proca theories have been generalized to include explicit gravitational interactions of a massive vector field [115–118]. The vector field Lagrangian is built so that precisely one extra (longitudinal) scalar mode propagates in addition to the two usual Maxwell-like transverse polarizations. Its full generalization contains terms with direct couplings between the vector and space-time curvature, whose structure closely resembles those of Horndeski's theory (A.1.25,A.1.26). In analogy to beyond Horndeski, there are also beyond generalized Proca interactions [119, 120]. Further extensions to multiple vector fields known as generalized multi-Proca/Yang-Mills theories are able to incorporate new couplings [121] and configurations [122], e.g. the extended triad $A_\mu^a = \phi^a \delta_\mu^0 + A(t)\delta_\mu^a$, as do theories with a vector and a scalar (Scalar-Vector-Tensor) [123]. For more details about these theories we recommend Ref. [61].

An iconic theory containing a vector is the Tensor-Vector-Scalar (TeVeS) theory by Bekenstein [124]. TeVeS emerged as a relativistic theory able to describe Modified Newtonian Dynamics (MOND), and thus as an alternative to dark matter. For an overview of field-theoretical aspects of TeVeS and related theories, including other relativistic MOND candidates, see Ref. [125]. TeVeS theory introduces several non-minimal ingredients. In addition to the gravitational metric $\tilde{g}_{\mu\nu}$ matter is minimally coupled to an effective metric

$$g_{\mu\nu} = e^{-2\phi}\tilde{g}_{\mu\nu} - 2\sinh(2\phi)A_\mu A_\nu \,, \tag{1.1.13}$$

which generalizes the scalar disformal relation (1.1.9), incorporating the vector. Here $\tilde{g}_{\mu\nu}$ is the gravitational metric, ϕ is the scalar. The vector A_μ is enforced to be time-like and normalized with respect to the gravitational metric $\tilde{g}^{\mu\nu}A_\mu A_\nu = -1$. TeVeS has a very rich phenomenology, including effects in GW propagation [126]. At the level of cosmology it is partially able to mimic DM, although the oscillations of the fields make it hard for the theory to reproduce the peaks in the CMB [127–129].

Massive gravity and tensor fields

Giving a mass to the graviton is another means to extend GR, with gravity mediated by a particle with mass m_g, spin $s = 2$ and $2s + 1 = 5$ polarization states (see [130] for bounds on the graviton mass). Weinberg theorem on the structure of GR relies on the infrared properties of the interactions: a mass term changes this structure. Despite this clear loophole, constructing a self-consistent theory of massive gravity, free of pathologies and with the right number of degrees of freedom proved an extremely hard endeavor that took nearly 70 years to complete. The linear theory of massive gravity was formulated in 1939 by Fierz & Pauli [131] as linearized GR plus a mass term

$$S_{\rm FP} = \int d^4x \, m_g^2 \left(h^{\mu\nu}h_{\mu\nu} - (\eta^{\mu\nu}h_{\mu\nu})^2 \right) . \tag{1.1.14}$$

It was later found that Fierz-Pauli theory was discontinuous and gave different results from GR in the limit $m_g \to 0$ (vDVZ discontinuity) [132, 133]. The discrepancy is due to the longitudinal polarization of the graviton (the helicity-zero mode) not decoupling

in that limit. Considering non-linear interactions solved the apparent discontinuity by hiding the helicity-zero mode, which is strongly coupled in regions surrounding massive bodies and effectively decouples, recovering the GR predictions when $m_g \to 0$ [134]. Despite this progress, massive gravity had another important flaw: all theories seemed to have an additional mode (known as Boulware-Deser (BD) ghost) that renders the theory unstable [135, 136].

Ghost Free Massive Gravity The apparent difficulties were overcome in de Rham-Gabadadze-Tolley theory (dRGT) [137], also known as ghost-free massive gravity (for current reviews on the theory see [138, 139]). dRGT is a ghost free theory propagating the 5 polarizations corresponding to a spin-2 massive particle, universally coupled to the energy-momentum tensor of matter (cf. Fig. 2.2). The ghost-free property was initially shown in the decoupling limit (in which the helicity-0 mode decouples from the other polarizations) and then in the full theory [140, 141] . The phenomenological deviations induced by massive gravity are primarily due to the helicity-0 mode. On small enough scales the Vainshtein mechanism [134] (see [142] for a review) effectively suppresses these interactions, leading to predictions very similar to GR on Solar System scales (however, new classes of solutions for black holes do exist, in addition to the usual ones [143]).

Massive gravity may offer a solution to the accelerating universe. A heuristic argument is that the force mediated by the massive graviton has a finite range $V \sim \frac{1}{r}\exp(-r/\lambda_g)$, weakening over distances larger than the Compton wavelength of the graviton $r \gtrsim \lambda_g = \hbar/(m_g c^2)$. Hence, if the mass of the graviton is $m_g \sim H_0$ then gravity weakens at late times and on cosmological scales, causing an acceleration of the cosmic expansion relative to the GR prediction. The program to apply massive gravity as a dark energy model has hit important barriers, as flat FLRW solutions do not exist in this theory [144]. Accelerating solutions without a cosmological constant (CC) do exist with open spatial hypersurfaces [145], but they are unstable [146]. Proposed solutions include the graviton mass being generated by the vacuum expectation value of a scalar [144] or deformations of the theory in which the BD ghost is introduced, which provides dynamical accelerating, but meta-stable solutions [147]. Alternatively, one could promote the coefficients of the potential to be functions of the Stueckelberg fields [148]. Other ways to make massive gravity dynamical include the addition of a new field, such as a scalar field, e.g. quasi-dilaton [149], or one (or several) dynamical tensors in bigravity (and multigravity).

Bigravity and Multigravity In order to write a mass term for the metric, dRGT incorporates an additional, non-dynamical tensor, akin to the occurrence of $\eta_{\mu\nu}$ in eq. (1.1.14). Massive gravity can be extended by including a kinetic term to the auxiliary metric, which becomes fully dynamical. This leads to the theory of bigravity (or bimetric gravity) [150], which contains two spin-2 particles: one massive and one massless. The same procedure can be extended to more than two interacting metrics, leading to multigravity theories [151]. In these constructions there is always one massless excitation of the metric (a combination of the different tensor fields), with all other excitations being massive.

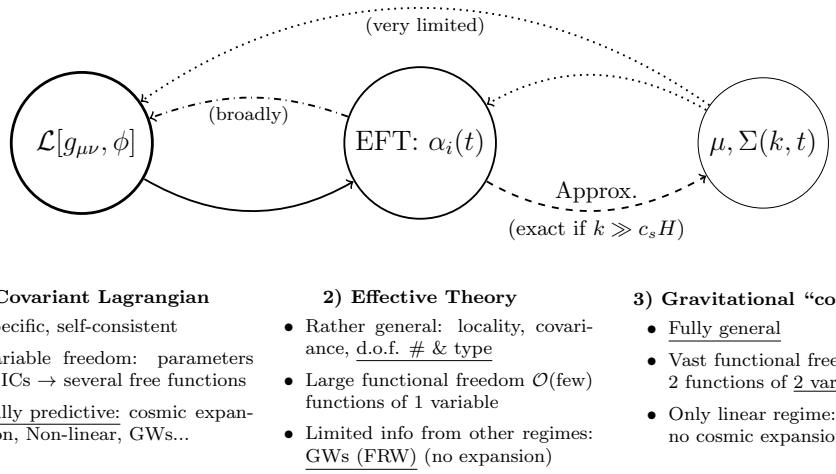

Figure 1.3. Effective descriptions of cosmological gravity, their relations and main advantages/shortcomings. Theories of gravity based on a gravitational Lagrangian are described in Sec. 1.1.1. The effective theory approach is described in 1.1.2 and the Gravitational "constants" in section 1.1.2.

Bigravity solves the problem of cosmological evolution, at least at the background level. Flat FLRW solutions do exist, and many viable expansion histories have been found to be compatible with data [152] and satisfying the Higuchi stability bound [153]. However, it was later found that these models had instabilities that affected the growth of linear perturbations [154], which were found to be quite generic across different branches of solutions [155]. In some cases the instabilities affect only scales sufficiently small for non-linear effects to be important (i.e. the Vainshtein mechanism) which might render the theory stable [156]. Another solution is to choose the parameters of the theory so instabilities occur at early times, when characteristic energies are high and bigravity is not a valid effective field theory. This happens by choosing a large hierarchy between the two Planck masses: the so-obtained theory is practically indistinguishable from GR plus a (technically natural) CC [157].

1.1.2 Descriptions of cosmological gravity

The immense variety of alternative theories has motivated the search for effective descriptions able to capture the phenomenology of generic dark energy models. The covariant actions approach reviewed in Sec. 1.1.1 offers several advantages, including 1) full predictivity, as (classical) solutions can be found from microscopic scales, to strong gravity and all the way to cosmology, 2) self-consistency, as different regimes can be computed for the same theory, leading to tighter constraints when the data is combined. For instance, following this approach, we discuss the cosmology of covariant Galileons in Sec. 2.2.1. Nonetheless, a great downside of this approach is that the predictions for every model/theory have to be obtained from scratch, which makes the exploration of the

theory space a daunting task.

An alternative route is to constrain deviations from GR, without reference to any fundamental theory. The tradeoff is to keep the theory of gravity as general as possible at the expense of dealing with a very simple space-time. The simplest situation is where the background space-time is flat and maximally symmetric (Minkowski), a setup useful to model gravity in the Solar System. In this simple case one can define a series of quantities, known as Parameterized Post-Newtonian (PPN) coefficients, that describe general modifications of gravity over Minkowski space (see Ref. [17] for details, including constraints and additional assumptions). These PPN parameters that can be constrained by experiments (such as the deflection of light by massive bodies) and computed for any theory, and thus provide a very efficient phenomenological dictionary.

In cosmology we are interested in describing gravity over a slightly less symmetric background: a spatially homogeneous and isotropic, but time evolving, Friedmann-Lemaitre-Robertson-Walker (FLRW) metric:

$$ds^2 = -(1+2\Psi)dt^2 + a^2(t)\left\{(1-2\Phi)\delta_{ij} + h_{ij}\right\}dx^i dx^j\,, \tag{1.1.15}$$

where metric perturbations are in Newtonian gauge with the sign conventions of Ma & Bertschinger [158]. The tensor perturbation is symmetric, transverse and traceless ($\partial^i h_{ij}, \delta^{ij} h_{ij} = 0$) and we have ignored vector perturbations. The time-evolution of the cosmological background makes an extension of PPN approach to cosmology a difficult task, as instead of constant coefficients one needs to deal with functions of time due to the evolution of the universe.

The most important example of an effective description in cosmology is the parameterization of the cosmological background, often done in terms of the equation of state $w \equiv p/\rho$ [159, 160]. Instead of computing the modifications to the Friedmann equations and the pressure and energy density contributed by the additional fields, a general approach to cosmological expansion is to specify $w(z)$ so that

$$H^2 = \frac{8\pi G}{3}(\rho_M + \rho_{\rm DE})\,, \tag{1.1.16}$$

$$\rho_{\rm DE} = \exp\left(-3\int d\log(a)(1+w)\right)\,. \tag{1.1.17}$$

This is sufficient to describe any cosmological expansion history and in any theory (as long as matter is minimally coupled and for perfect fluids) just by using the Friedmann equation (1.1.16) as a definition for $\rho_{\rm DE}$.

Describing the perturbations requires more functional freedom. Here we will review two common procedures, namely the effective theory of dark energy and the modified gravitational "constants". The different approaches (including the covariant theory approach), their features and connections are outlined in Fig. 1.3. Consistency checks between the background and perturbations can also be used to test the underlying gravity theory [161, 162].

Effective theory of Dark Energy

The effective (field) theory of dark energy (EFT-DE) [163–165] can be used to systematically describe general theories of gravity over a cosmological background (see Ref. [166] for a review). The original formulation applies to theories with a scalar field ϕ and uses the unitary "gauge": a redefinition of the time coordinate as the constant ϕ hypersurfaces (this is always possible if $\phi_{,\mu}$ is time-like and non-degenerate, as in perturbed cosmological backgrounds, but not in general). One then constructs all the operators compatible with the symmetries of the background (recalling that the time translation invariance is broken by the cosmological evolution).

A very convenient basis for the EFT functions was proposed by Bellini & Sawicki [167], when restricted to Horndeski's theory. In their approach the EFT functions are defined by the kinetic term of the propagating degrees of freedom in the equations of motion. The dynamical equation for tensor perturbations

$$\ddot{h}_{ij} + (3 + \alpha_M)\dot{h}_{ij} + (1 + \alpha_T)\frac{k^2}{a^2}h_{ij} = 0\,, \tag{1.1.18}$$

introduces two dimensionless functions

- *tensor speed excess* α_T describes the modification in the GW propagation speed $c_g^2 = (1 + \alpha_T)$. This modification is frequency independent (see chapter 3).
- *Planck-mass run rate* α_M enters as a friction term. It is related to the *cosmological strength of gravity* M_*^2 (the kinetic term of tensor perturbations) by $\alpha_M = \frac{d\log\left(M_*^2\right)}{d\log a}$ (see chapter 4).

The equations in the scalar sector (eqs. (3.20), (3.21) of [167]) can be used to define the remaining functions. If we look only at the second time derivatives (that is, the kinetic terms)

$$2\ddot{\Phi} - \alpha_B H\delta\ddot{\phi}/\dot{\phi} + \quad \cdots \quad = 0\,, \quad \text{(ii-trace)} \tag{1.1.19}$$

$$\alpha_K\delta\ddot{\phi}/\dot{\phi} + 3\alpha_B\ddot{\Phi}/H + \quad \cdots \quad = 0\,, \quad (\phi \text{ scalar}) \tag{1.1.20}$$

(note the ellipsis denote terms without second time derivatives) one can define

- *braiding*, or *kinetic gravity brading* α_B quantifies mixing between the second derivatives of the metric in the field equation (and vice versa). This is a generic property of modified gravity [87, 168].
- *kineticity* α_K modulates the "stiffness" of the scalar field (how hard it is to excite perturbations in ϕ). The kineticity is intimately related to the propagation speed of scalar perturbations, which satisfies $c_s^2 \propto \alpha_K^{-1}$: higher kineticity values lead to slower scalar waves and vice versa.

These functions can be computed from the Lagrangian functions in (1.1.4), and for a given theory will depend on the value of the scalar field and its time derivative. Constraints on the α-functions can also be used to reconstruct the terms in a fundamental

	Horndeski					DHOST		
	$G_{2,\phi}$	$G_{2,X}$	$G_{3,X}$	$G_{4,\phi}$	$G_{4,X}\cdots$	GLPV	$\mathcal{C}_1$	$\mathcal{C}_2$
$1+w$	✓	✓	✓	✓	✓	✓	✓	✓
α_K	–	✓	✓	✓	✓	✓	✓	✓
α_B	–	–	✓	✓	✓	✓	✓	✓
α_M	–	–	–	✓	✓	✓	✓	✓
α_T	–	–	–	–	✓	✓	✓	✓
α_H	–	–	–	–	–	✓	✓	✓
β_1	–	–	–	–	–	–	✓	•
α_L	–	–	–	–	–	–	–	✓

– zero, ✓ non-zero (arbitrary), • non-zero (constrained)

Table 1.1. EFT functions in scalar-tensor theories: a hyerarchy exists by which more complex theories of gravity (left to right) produce a larger set of effects (more non-zero functions). For the DHOST theories there are two classes of degeneracy conditions: $\mathcal{C}_1$ and $\mathcal{C}_2$. Some non-trivial special cases are known to exist: $f(R)$ and $f(G)$ theories have $\alpha_K = 0$, while first generation theories (1.1.3) including $f(R)$, Brans-Dicke satisfy $\alpha_B + \alpha_M = 0$ [167] and 2 beyond Horndeski combinations produce $\alpha_T = 0$ [5, 99] (see section 3.4).

theory, as shown in Tab. 1.1. Systematic reconstructions of the Lagrangian from the α functions have been also explored [169, 170].

Increasingly complex theories of gravity lead to a larger number of EFT functions. In beyond Horndeski theories of the GLPV type, e.g. (1.1.11,1.1.12), a new function α_H is introduced [101] which phenomenologically produces a weakening of gravity on small but linear cosmological scales [171]. In DHOST theories including (1.1.10) the situation is more involved, as the new EFT functions $(\alpha_L, \beta_1, \beta_2, \beta_3)$ need to be related to each other and α_T, α_H by the degeneracy conditions that prevent the introduction of additional degrees of freedom [108]. This leads to two classes of theories with one free function, which is either α_L or one among β_i. New EFT functions appear beyond scalar-tensor theories, as has been explicitly derived for vector-tensor [172] and bimetric [173] theories (including bimetric gravity), with a unified treatment of theories with different degrees of freedom [174].

Different versions of the linear EFT-DE approach has been implemented in numerical codes able to obtain predictions based on linear perturbation theory. Publicly available implementations exist in EFTCAMB [175], `hi_class` [176] and COOP [177], with the first two based on the CAMB and CLASS Boltzmann codes [178, 179]. In addition, the CLASS-Gal code (integrated into CLASS) can be used to compute relativistic corrections to cosmological observables [180]. These and other codes have been tested against a large class of models at a level of precision sufficient for current and next-generation cosmological experiments [181].

The EFT framework has been tested using linear observables. Horndeski theories were tested against current experiments, leading to $\mathcal{O}(0.1-1)$ constraints on the α-functions varying over $\alpha_B, \alpha_M, \alpha_T$ [182], with $\alpha_M = -\alpha_B$ [183] and setting $\alpha_T = 0$ to reflect the strong bounds on the GW speed [184] (α_K is very weakly constrained by

current data). Future experiments have great potential to improve on these bounds, and are expected to improve the sensitivity to $\mathcal{O}(0.01 - 0.1)$ [185–189]. EFT-based modifications of gravity might be observable through relativistic effects on ultra-large scales [187, 190] (see also the discussion in Sec. 1.1.2): these techniques might improve significantly our ability to constrain α_K, although it will remain the hardest to measure [186]. Those works used simple functional dependence of the EFT functions. It has been nonetheless shown that simple parameterizations are indistinguishable from more complex models in most cases, even for next-generation cosmology experiments [191].

The EFT approach has been generalized beyond linear perturbations for Horndeski theories. Including non-linear cosmological perturbations in general introduces new functions at every order in perturbation theory (e.g. to compute the bispectrum [192]). However, a restriction to cubic and quartic operators (in the unitary gauge) leads to only 3 new operators on quasi-static scales [193]. Some applications of non-linear EFT-DE include corrections to the power spectrum (e.g. [194]), the use of higher-order correlations as a test of gravity, such as the bispectrum of matter [192], galaxies [195] and CMB lensing [196] or the the non-linear shift of the BAO scale [197].

Modified Gravitational "constants"

A very commonly used approach employs general modifications of the equations relating the gravitational potentials to the matter density contrast

$$\nabla^2\Psi = 4\pi G a^2 \mu(t,k)\rho\delta\,, \tag{1.1.21}$$

$$\nabla^2(\Phi+\Psi) = 8\pi G a^2 \Sigma(t,k)\rho\delta \tag{1.1.22}$$

(note that different conventions exist in the literature). Here δ is the density contrast in the Newtonian gauge (1.1.15) and the functions μ, Σ parameterize the evolution of the gravitational potentials as a function of time a and scale k. The functions μ, Σ are often referred to as $G_{\rm matter}$, $G_{\rm light}$ because gradients of Ψ determines the force felt by non-relativistic particles and those of $\Psi+\Phi$ the geodesics of massless particles (and thus the lensing potential). The ratio of the gravitational potentials,

$$\eta \equiv \frac{\Phi}{\Psi} = \frac{2\Sigma}{\mu} - 1\,, \tag{1.1.23}$$

is of particular interest, since GR predicts that it is exactly one in the absence of radiation and any sizable deviation could be an indication of modified gravity.

This approach has numerous advantages as a test of gravity against data. It is completely theory agnostic, not requiring any information on the ingredients or laws of the theories being tested. Most importantly, it is completely general for universally coupled theories: given any solution $\Delta, \Psi, \Phi(a,k)$ it is possible to obtain μ, Σ through (1.1.21,1.1.22). In this sense, any finding of $\mu, \Sigma \neq 1$ might point towards deviations from GR and warrant further investigation.

The main shortcoming of this approach is its great generality: any practical attempt to implement (1.1.21,1.1.22) requires a discretization of the functional space, introducing $2 \cdot N_k \cdot N_z$ free parameters for a homogeneous binning. In contrast, the EFT approach for Horndeski theories (1.1.18,1.1.19) requires only $4 \cdot N_z$ parameters, making it a

more economic parameterization for all but the simplest scale-dependencies ($N_k = 1, 2$). Capturing the full scale dependence of μ, Σ requires either a large parameter space or assumptions about the k-dependence.

A common practice to overcome this limitation is to choose a functional form for μ, Σ as a function of scale. For Horndeski theories the functional form is a ratio of quadratic polynomials in k [198, 199]

$$\mu = h_1 \frac{1 + h_5 k^2}{1 + h_3 k^2}\,, \quad \eta = h_2 \frac{1 + h_4 k^2}{1 + h_5 k^2}\,, \tag{1.1.24}$$

for functions h_i that depend on redshift through the theory (1.1.4) and the scalar field evolution. The mapping is exact on small scales in which the field dynamics can be neglected, below the scalar sound horizon [200]. A k-dependence as the ratio of polynomials is generic in local theories at quasi-static scales [201], with higher order polynomials possible in Lorentz-violating [202], multi-field [203] theories. Studies with current data have tested rather simple parameterizations of μ, Σ: for instance the Planck survey tested the case of k-independent μ, η in addition to the theory-motivated (1.1.24) [183]. Future surveys will improve the resolution on the scale-dependence: 3 k-bins are the minimum to constraint all the parameters in eq. (1.1.24), with 6 bins in z [204, 205]. A limited handle on scale-dependence on ultra-large scales might be achievable [206, 207] (see also [208–210] for related parametrizations).

Another main shortcoming of the completely general approach is that there is no information from other regimes. The major setback with respect to EFT is the lack of information from gravitational wave observables, while in EFT the tensor and scalar sectors are modified accordingly i.e. GW data restrict the modifications available to scalar perturbations, for instance, theories with $\eta \neq 1$ require either α_M or α_T to be non-zero [211]. Attempts to explore the connections between μ, Σ and the EFT approach in Horndeski-like theories have used very general parameterizations: connecting theoretical viability conditions of the theory with the behavior of μ, η [212], including the case with $\alpha_T = 0$ to address the impact of the GW speed measurement [213]. General properties of Horndeski theories could be inferred from detailed measurements of μ, Σ [214]. Similarly to the EFT approach, the background evolution is unknown and the equation of state (1.1.17) is in principle arbitrary. However, theoretical priors on $w(z)$ can be obtained for broad classes of Lagrangians (e.g. quintessence [215]) or from stability conditions in general realizations of the EFT functions [216].

1.2 Dark matter and the early universe

Dark matter is the other main unknown in the Standard Model of Cosmology. Over the decades, there has been accumulating evidence of the presence of more matter in the universe that the one we see with our telescopes: from galaxy rotation curves, to gravitational lensing, to the cosmic microwave background [217]. Still, we lack of a fundamental understanding of this new component. Within the SM of Cosmology, DM is assumed to be a non-baryonic, pressure-less perfect fluid where cosmic structure form

in a hierarchical manner (from small structures to large ones). This model successfully describe the growth of cosmological perturbations from the small inhomogeneities seeded by inflation to the formation of galaxies and clusters of galaxies. This is unquestionably a remarkable achievement, capable of explaining five orders of magnitude of growth in the perturbations over 13 billion years. However, the large scale success might be challenged be long standing small scale problems such as the missing satellites or the core versus cusp issue [218]. Moreover, the observation of high redshift active galaxy nuclei questions the hierarchical formation mechanism of super massive BHs (SMBHs) [219]. This could be either an indication of a fracture in the cold DM paradigm, a sign of the difficulty of including baryonic physics in numerical simulations, or just that we do not have good enough tests at scales $\lesssim 10$kpc.

Before moving on, we should remember that the cold DM scenario is also based on the validity of GR. Since present evidence of DM is all derived from gravitational observables, it is natural to ask wether gravity could be the problem to begin with. However, alternative gravity theories have serious difficulties to explain DM phenomena at different scales from galaxies, to clusters of galaxies, to the CMB. This is the case of TeVeS, the covariant theory of modified Newtonian dynamics that we have introduced in section 1.1.1. It can explain the rotation of galaxies but fails to explain the peaks in the CMB [127–129].

1.2.1 The origin of DM

DM has a fundamental role throughout the history of the universe. Big Bang nucleosynthesis provides the earliest information of the amount of baryonic matter, implying that DM should be present at least before the temperature cooled down to $T \sim 10$MeV. Therefore, the origin of DM should be in the early universe.

The standard paradigm in cosmology assumes DM to be made of weakly interacting massive particles (WIMPs) [220]. The main advantage of this model is that if they interact via weak interactions, the relic abundance automatically has the right order of magnitude. The down side is that this DM candidate has not been detected yet, despite of the substantial efforts with both direct and indirect searches. Another popular candidate are ultra-light particles [221] with masses of the order of $m \sim 10^{-22}$eV and Compton wavelength at the kpc scale (also known as fuzzy DM). The interest in this model resides in its promise to solve some of the small scale problems.

Although DM is mostly thought as an undiscovered elementary particle, it might have a much less exotic origin (in the sense of physics beyond the SM or not). DM might be composed of massive compact objects whose radiation cannot be detected, such as neutron stars or black-holes, generically named MAssive Compact galactic Halo Objects (MACHOs). In particular, black-holes stands, a priori, as natural DM candidates since

i) they do not emit light and interact only gravitationally,

ii) they move at non-relativistic speeds $v/c \ll 1$,

iii) they are basically collision-less due to their small size, with an interacting cross-section

$$\sigma_{\rm Sch} \sim 3 \cdot 10^{-11} \times \left(\frac{M_{\rm BH}}{10^{22}{\rm g}}\right)^2 {\rm cm}^2 \,. \tag{1.2.1}$$

Of course, these BHs could not be of astrophysical origin, since they might be present in the universe in the first fractions of a second.

The idea that BHs could have a primordial origin dates back to Hawking [222]. Soon after, it was proposed that this primordial BHs (PBHs) could be the seed of galaxies and the cosmic structure [223]. Interestingly, if these PBHs are heavy, $10^2 - 10^3 M_\odot$, they might explain the formation of SMBHs [224, 225]. As of today, there is a plethora of mechanisms to form PBHs [226]. In this thesis, we will focus on PBHs formed from large curvature fluctuations generated during inflation. We will present this formation mechanism in chapter 6. For the moment, let us concentrate on how to probe this DM proposal.

1.2.2 The quest of the compactness of DM

Provided that there is a mechanism to form PBHs in the early universe and in the right amount, they would be basically indistinguishable from particle DM at large scales. The key question is then how to distinguish both scenarios, fluid DM versus compact DM.

There are four main channels in which PBH DM could be probed:

- *Gravitational lensing:* the best way to asses the compactness of DM is to directly measure the gravitational potential. This can be achieved with gravitational lensing. Lensing techniques can be applied to sources and lenses of various type. For this particular problem, the goal is to detect the magnification of the source's light when the PBH crosses the line of sight between the source and the observer. The duration of the amplifications scales with the mass of the lens. More massive lenses require longer monitoring campaigns. There are three principal types of lensing searching for compact objects. The first type is *microlensing* [227], monitoring stars in the Magellanic Clouds. Collaborations like MACHO [228], EROS [229] and OGLE [230] have placed upper bounds on the amount of compact objects in the $10^{-3} - 1 M_\odot$ range. Much smaller masses, $10^{-11} - 10^{-5} M_\odot$, have been probed with the Hyper Suprime-Cam that have observed tens of million stars in M31 [231] and Kepler [232]. There is also *femtolensing* which probes BHs in the range $10^{-16} - 10^{-14} M_\odot$ by looking for an oscillatory pattern in the photon spectrum of gamma ray bursts [233, 234]. Finally, one could search for *supernova lensing* [235], which is sensitive to compact objects with masses larger than $10^{-2} M_\odot$.

- *Dynamical effects:* the presence of compact objects also affects astrophysical systems by gravitational interactions. In particular, BHs could disrupt white dwarfs, neutron stars, wide binaries or globular clusters [236]. Therefore, by observing such astrophysical object one could infer the abundance of PBHs. However, this class of tests usually suffer from large astrophysical uncertainties that make them

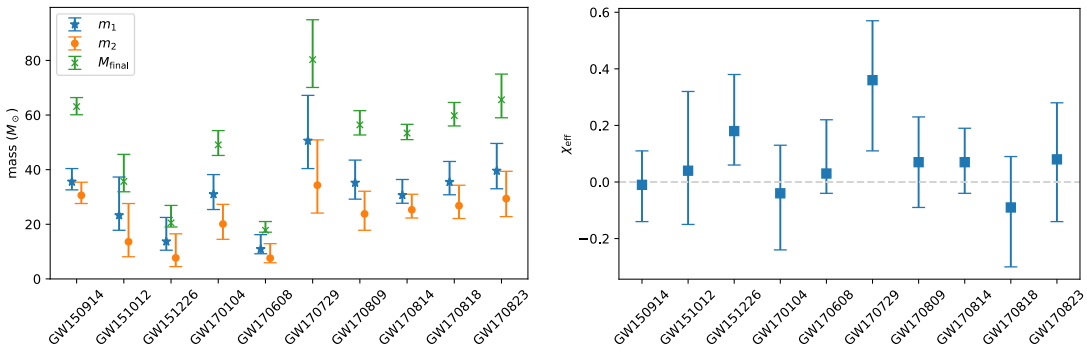

Figure 1.4. Masses (of each individual BH m_1, m_2 and final $M_{\rm final}$) and effective spin $\chi_{\rm eff}$ for the 10 binary black-holes reported by the LIGO-Virgo collaboration in runs O1-2 [24] presented in the left and right panels respectively.

unreliable. Similarly, the presence of PBHs in dwarf galaxies would make the stars to move faster and spread wider. Observations of star clusters in the dwarf galaxy Eridanus II may constrain masses larger than $100M_\odot$ [237].

- *Gravitational waves:* if PBH form binaries that coalesce within the age of the universe, they would produce GWs that could be detected individually or in stochastic backgrounds. PBH binaries could form in the early universe during radiation domination [238]. Although at formation the mean separation between PBHs is larger than the Hubble horizon, as the universe expands PBHs could find each other and form bound systems. Once a binary is formed, it would emit GWs until merger. PBH binaries could also form in the present universe by close encounters in DM halos. The merger rate of binaries formed during radiation domination [239] is larger than those formed in present times [240]. A clustered PBH distribution with a broad range of masses also allows to form binaries in the early universe [241]. The formation mechanism of BHs detected with GWs can be constrained from their mass spectrum, spin distribution, eccentricities and merger rate as a function of redshift.

  We summarize the masses and effective spins of the 10 BBHs detected in LIGO-Virgo runs O1-2 in Fig. 1.4.[6] So far, BBHs with masses $\mathcal{O}(30M_\odot)$ and mass ratios $q = m_2/m_1 \gtrsim 1/2$ have been observed to form BHs of $\mathcal{O}(50M_\odot)$. Due to these (unexpected) large masses, already with the first detection the question of the possible primordial origin was raised [239–241]. Later analysis have shown also plausible astrophysical formation mechanisms, see e.g. [244]. Their effective spins $\chi_{\rm eff}$, which is the mass weighted projection of the two (dimensionless) spins of the

[6]Note that an independent reanalysis of LIGO data by an external group recently found one more event in O1 [242] and six more (although the credibility of three of them is low) in O2 [243]. Moreover, as pointed out in footnote 1, there are new candidates in the ongoing O3.

binaries $\vec{\chi}_i$ onto the Newtonian orbital angular momentum $\vec{L}$

$$\chi_{\rm eff} = \frac{m_1 \cdot \vec{\chi}_1 + m_2 \cdot \vec{\chi}_2}{m_1 + m_2} \cdot \vec{L}\,, \tag{1.2.2}$$

are found to be small, close to 0. This result challenges the standard stellar evolution formation channels and might indicate that the BHs forming the binary are born with low spin ($\chi_i \sim 0$) [245]. PBHs formed during radiation domination are expected to be form with low spin [246].

In section 7.2.2 we will discuss in more detail the GW signatures of PBH scenarios, using as a proxy a critical Higgs inflation model.

- *Energy injections:* BHs formed in the early universe leave several imprints along the history of the universe due to the energy injections that they transfer to the cosmic medium. Specifically, constraints can be derived in the abundance of massive PBH, $M > 10^2 M_\odot$, from the temperature and polarization anisotropies of the CMB that limit the accretion of PBHs [247]. Also, one can limit the abundance of microscopic PBHs from the absence of a extragalactic γ-background produced by Hawking evaporation [248].

Altogether, compact DM scenarios can be probed by several means. In Fig. 7.6 we summarize constraints on the abundance of monochromatic distributions of PBHs, represented by the fraction of the DM that they can constitute $\Omega_{\rm PBH}/\Omega_{\rm DM}$.

CHAPTER

2

Basics of gravitational waves

Gravity is a universal, long-range force. This, in field theory language, implies that it must be described by a metric field $g_{\mu\nu}$ in order to manifestly preserve locality and Lorentz invariance. At low energies, the leading derivative interactions are second order. Therefore, gravity theories generically predict the existence of propagating perturbations or, in other words, the existence of GWs. One can define a metric perturbation $h_{\mu\nu}$ as a small difference between the metric field $g_{\mu\nu}$ and the background metric $g^{\rm B}_{\mu\nu}$

$$h_{\mu\nu} = g_{\mu\nu} - g^{\rm B}_{\mu\nu}\,, \tag{2.0.1}$$

where $|h_{\mu\nu}| \ll 1$. However, in curved space it is non-trivial to distinguish the perturbation from the background unless the latter possesses some degree of symmetry, e.g. flat space or FLRW. A way out is to define GWs via geometric optics [249]. In this context, the key element to distinguish the GW from the background is the size of the fluctuations $\lambda_{\rm gw}$ with respect to the typical size of the background variation $L_{\rm B}$. One could associate the typical variation scale in the background with the minimum value of the components of the background Riemann tensor

$$L_{\rm B} \sim |R^{\rm B}_{\alpha\beta\gamma\rho}|^{-1/2}\,. \tag{2.0.2}$$

For astrophysical sources, we will see later that the wavelength of the GW $\lambda_{\rm gw}$ is orders of magnitude smaller than the typical variations of $L_{\rm B}$ for cosmological setups. The fact that $\lambda_{\rm gw} \ll L_{\rm B}$ implies that there is a clear hierarchy between background and perturbations, allowing to solve the problem using an adiabatic (or WKB) expansion.

In the following, we describe the basics of GWs. We begin by introducing GWs in GR. Then, we explore the propagation in cosmological backgrounds. Subsequently, we describe how this picture is changed beyond GR. Finally, we discuss the status of present and future GW detectors. We recommend the reader Ref. [249–253] for more details.

2.1 GWs in General Relativity

General Relativity is a universal, infinite-range force. As we have seen in the previous section, this implies that it is described by a massless, spin-2 field. The dynamics is described by Einstein's equations (1.1.2). Importantly, not all the components of the Einstein tensor $G_{\mu\nu}$ contain second order time derivatives of the metric $g_{\mu\nu}$. This implies

that not all of the 10 components of $g_{\mu\nu}$ will propagate. In particular, the $G_{0\mu}$ equations act as 4 constraint equations. This, together with the 4 unphysical modes reduced by the gauge choice, leaves only 2 propagating degrees of freedom. This is precisely what one would expect for a massless spin-2 particle.

In order to study GWs, the next step is to study the linearized Einstein's equations. To diagonalize the equations for the tensor perturbations, one has to introduce the trace-reversed perturbation

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} h g^{\rm B}_{\mu\nu} \,, \tag{2.1.1}$$

whose name comes from the fact that $\bar{h} = -h$ where $h = g^{\mu\nu}_{\rm B} h_{\mu\nu}$ and $\bar{h} = g^{\mu\nu}_{\rm B} \bar{h}_{\mu\nu}$ are the traces of $h_{\mu\nu}$ and $\bar{h}_{\mu\nu}$ respectively. Fixing the Lorenz gauge for this new variable $\nabla^\mu \bar{h}_{\mu\nu} = 0$, the linearized Einstein equations in curved space-time read

$$\begin{aligned} &\Box \bar{h}_{\mu\nu} + 2 R^{\rm B}_{\mu\alpha\nu\beta} \bar{h}^{\alpha\beta} = \\ &-16\pi G \delta T_{\mu\nu} + 2 R^{\rm B}_{(\mu}{}^{\alpha} \bar{h}_{\nu)\alpha} - R^{\rm B} h_{\mu\nu} + g^{\rm B}_{\mu\nu} R^{\alpha\beta}_{\rm B} \bar{h}_{\alpha\beta} \,, \end{aligned} \tag{2.1.2}$$

where covariant derivatives are built with the background metric $g^{\rm B}_{\mu\nu}$. Here, we have introduced the perturbed energy-momentum tensor $\delta T_{\mu\nu}$ as the difference of the total energy momentum tensor $T_{\mu\nu}$ with respect to the background solution $8\pi G T^{\rm B}_{\mu\nu} = R^{\rm B}_{\mu\nu} - \frac{1}{2} g^{\rm B}_{\mu\nu} R^{\rm B}$. One should note that, in vacuum, all the Ricci tensors vanish in the second line. Moreover, for short-wave GWs $\lambda_{\rm gw} \ll L_{\rm B}$, the Riemann tensor in the first line has a subdominant contribution.

To deal with the two GW polarizations, it is convenient to work in the transverse-traceless (TT) gauge, which is defined by

$$h_{0\mu} = 0 \,, \quad \nabla^j h_{ij} = h^i{}_i = 0 \,. \tag{2.1.3}$$

Note that in the TT gauge, the trace-reversed perturbation $\bar{h}_{\mu\nu}$ is equal to the original perturbation $h_{\mu\nu}$. If the GW is propagating in the z-direction, the spatial components become

$$h_{ij} = \begin{pmatrix} h_+ & h_\times & 0 \\ h_\times & -h_+ & 0 \\ 0 & 0 & 0 \end{pmatrix} \,, \tag{2.1.4}$$

with h_+ and $h_\times$ being the two polarizations of GR.

2.1.1 Generation

A first question to address is how GWs are produced. Let us consider a GW source in vacuum within the short-wave approximation. Then, the general propagation equation (2.1.2) reduces to

$$\Box \bar{h}_{\mu\nu} = -16\pi G \, \delta T_{\mu\nu} \,. \tag{2.1.5}$$

This wave equation can be solved in analogy to electromagnetism using a Green's function. In terms of the retarded time $t_r = t - |\vec{x} - \vec{y}|$, the solution is

$$\bar{h}_{\mu\nu}(t, \vec{x}) = 2G \int d^3\vec{y} \, \frac{\delta T_{\mu\nu}(t_r, \vec{y})}{|\vec{x} - \vec{y}|} \,. \tag{2.1.6}$$

For an isolated, far away, non-relativistic source, this solution can be simplified. In fact, one can make a multipole expansion. The zeroth moment corresponds to the mass-energy of the source $M = \int T^{00}(t, \vec{y}) d^3\vec{y}$. However, conservation of energy for an isolated source tells us that M cannot vary in time. Next, the mass dipole moment $M_i(t) = \int y_i T^{00}(t, \vec{y}) d^3\vec{y}$ is associated to the motion of the center of mass. Nevertheless, its time derivative is the momentum of the source that also has to be conserved[1]. Consequently, the leading contribution is the mass quadrupole moment $M_{ij}(t) = \int y_i y_j T^{00}(t, \vec{y}) d^3\vec{y}$, which generates GWs through its second time derivatives

$$\bar{h}_{ij}(t, \vec{x}) = \frac{2G}{r} \frac{d^2 M_{ij}}{dt^2}(t_r) \,. \tag{2.1.7}$$

For a binary system of masses m_1 and m_2, the quadrupole radiation is

$$h_{+,\times} = \frac{\mathcal{M}_c^{5/3} f^{2/3}}{r} F_{+,\times}(\text{angles}) \cos \Phi(t) \,, \tag{2.1.8}$$

where F is a function of the orientation of the binary that depends on the polarization $+$ or $\times$ (recall (2.1.4)), $\Phi(t)$ is the phase and we have introduced the chirp mass

$$\mathcal{M}_c = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}} \,. \tag{2.1.9}$$

As the masses orbit one around the other, they will lose energy with the emission of GWs. They will begin getting closer and orbiting faster until they eventually merge. Thus, the frequency of GWs will increase with a characteristic chirp signal following

$$\dot{f}_{\text{gw}} = \frac{96}{5} \pi^{8/3} \left(\frac{G\mathcal{M}_c}{c^3} \right)^{5/3} f_{\text{gw}}^{11/3} \,. \tag{2.1.10}$$

Note that to consider the energy loss due to GWs emission one has to go to second order in perturbation theory. An example of the typical GW strain and frequency of a compact binary coalescence is presented in Fig. 2.1.

Typical binary compact objects emitting detectable GWs are binary neutron stars (BNS) and binary black-holes (BBH). The order of magnitude of the frequency of the GWs of these systems is

$$f_{\text{gw}} \sim \frac{1}{4\pi} \left(\frac{3GM}{R^3} \right)^{1/2} \sim 1\text{kHz} \left(\frac{10 M_\odot}{M} \right) \,, \tag{2.1.11}$$

where $M_\odot$ is equal to one solar mass. This implies that higher masses lead to lower frequencies. In terms of the wavelength one finds

$$\lambda_{\text{gw}} \sim 200\text{km} \left(\frac{M}{M_\odot} \right) \,. \tag{2.1.12}$$

[1]Similar arguments apply for the spin angular momentum in case the source exhibit some internal motion.

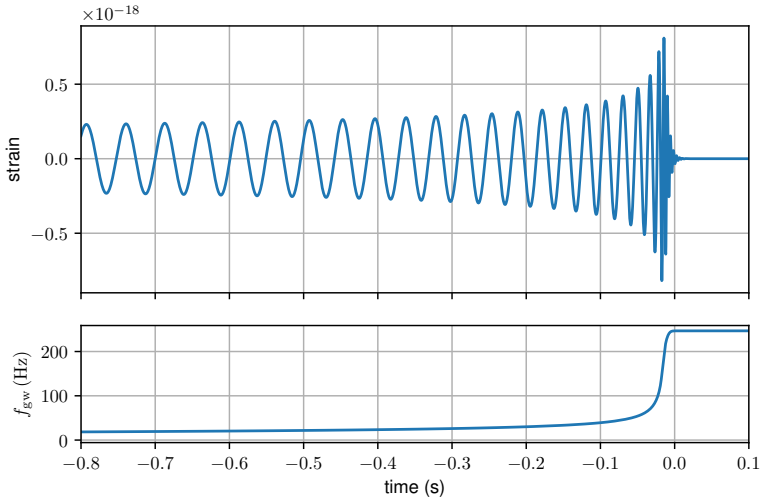

Figure 2.1. Typical GW signal of a compact binary coalescence. The GW strain (above) and the GW frequency (below) are plotted as function of the time before merging. This waveform is a template of the first event detected GW150914 [254].

This allows us to compare the size of the wavelength with the typical size of the background curvature variation $L_{\rm B}$. For cosmology, the size of the curvature is related to the Hubble horizon $L_{\rm B}^{\rm cosmo} \sim 10^{26}$m. For our galaxy one can estimate $L_{\rm B}^{\rm gal} \sim 10^{23}$m and for the Solar System $L_{\rm B}^{\rm SolSys} \sim 10^{16}$m. As it can be observed, the geometric optics expansion is an excellent approximation due to the great hierarchy between $\lambda_{\rm gw}$ and $L_{\rm B}$. Only GWs passing near a very dense object such as a BH, $L_{\rm B}^{\rm BH} \sim (M_{\rm BH}/M_\odot)$km, would break this short-wave approximation.

The typical amplitude of a GW from a compact binary can be estimated using (2.1.8), leading to

$$h \sim 10^{-21} \left(\frac{\mathcal{M}_c}{10\,M_\odot}\right)^{5/3} \left(\frac{f}{100\,\text{Hz}}\right)^{2/3} \left(\frac{100\,\text{Mpc}}{r}\right) . \tag{2.1.13}$$

Contrary to EM waves, GW detectors are directly sensitive to the amplitude of the wave, which falls like $1/r$ and not as the luminosity $1/r^2$. This means that even if the amplitudes are very small, GW detectors are more sensitive to distant sources.

2.1.2 Propagation

Once the GW is generated, it will propagate in vacuum following

$$\Box \bar{h}_{\mu\nu} + 2R^{\rm B}_{\mu\alpha\nu\beta}\bar{h}^{\alpha\beta} = 0\,. \tag{2.1.14}$$

A general solution of this wave equation can be written as the sum of plane waves

$$\bar{h}_{\mu\nu}(t,\vec{x}) = \text{Re}\left[A_{\mu\nu}\cdot e^{ix_\alpha k^\alpha}\right] , \tag{2.1.15}$$

where Re denotes the real part. By plugging this expression in the wave equation and expanding in powers of k, one finds at leading order that

$$k_\mu k^\mu = g_{\rm B}^{\mu\nu} k_\mu k_\nu = 0\,. \tag{2.1.16}$$

Therefore, GWs propagate in null geodesics determined by the background metric. This means that the GW-cone is the same as the light-cone and both waves propagate at the same speed. Moreover, the wave is transverse to the propagation direction

$$k^\mu A_{\mu\nu} = 0\,, \tag{2.1.17}$$

similarly to electromagnetic waves. Finally, by defining the scalar amplitude $\mathcal{A} = \left(\frac{1}{2} A^*_{\mu\nu} A^{\mu\nu}\right)^{1/2}$ one realizes that

$$\nabla_\alpha \left(k^\alpha \mathcal{A}\right) = 0\,, \tag{2.1.18}$$

which can be interpreted as the conservation of gravitons. One should note that $R^{\rm B}_{\mu\alpha\nu\beta}$ in the wave equation only modifies the amplitude at second order. Consequently, at first order in geometric-optics, the wave equation $\Box \bar{h}_{\mu\nu} = 0$ can be rewritten as

$$\Box R^{\rm gw}_{\mu\alpha\nu\beta} = 0\,. \tag{2.1.19}$$

This expression could be used as a gauge invariant, coordinate independent definition of the propagation of GWs in vacuum.

2.1.3 Detection

To see the effect of a GW passing by, one has to study the deviation of nearby geodesics. Given two particles with four-velocity U^μ separated by S^μ, their separation evolves as

$$\frac{D^2 S^\mu}{d\tau^2} \equiv U^\rho \nabla_\rho \left(U^\gamma \nabla_\gamma S^\mu\right) = R^\mu{}_{\alpha\beta\nu} U^\alpha U^\beta S^\nu\,, \tag{2.1.20}$$

where τ is the proper time. At leading order, the four velocity is just the unit vector $U^\mu = (1,0,0,0) + \mathcal{O}(h)$, and we only have to compute the Riemann tensor in the TT gauge. The result is

$$\frac{\partial^2 S^\mu}{\partial t^2} = \frac{1}{2} S^\nu \frac{\partial^2}{\partial t^2} h^\mu{}_\nu\,, \tag{2.1.21}$$

where we have also used that to leading order the proper time τ and the coordinate time t coincide. Accordingly, only the components of the separation vector S^μ transverse to the propagation vector will feel the effect of the GW. In these directions, the separation between the test particles will oscillate as the GW travels perpendicular to them. In Fig. 2.2, we plot the effect of the different GW polarizations crossing a circle of test masses.

GW detectors precisely rely on this principle that GWs can alter the separation between test masses. Modern detectors are interferometers. In brief, they are constituted by two perpendicular arms of the same length with two mirrors in free fall at their ends (acting as test particles). A laser beam is split in the two arms so that the beams

reflect in each mirror and come back to the splitting point. In the absence of a GW, both laser beams returning will interfere destructively and no signal would arrive to the detector. However, if a GW crosses the interferometer, it will change the length of the arms differently. This means that the laser beams will take different times to travel the arms, arriving at the splitting point with different phases. Then, the destructive interference is lost and some signal gets to the detector.

Note that the typical distance variation δL of two test masses separated by L is approximately $\delta L \sim h \cdot L$. For compact binaries, we have seen that the strain amplitude is $h \sim 10^{-21}$. Therefore, LIGO-type detector with arms of the order of kilometers have to measure distance variations

$$\delta L \sim 10^{-18} \left(\frac{h}{10^{-21}}\right) \left(\frac{L}{\mathrm{km}}\right) \mathrm{m}\,, \tag{2.1.22}$$

a thousand times smaller than the nucleus of an atom. To achieve that, each arm has a resonant cavity in which the laser beams bounce back and forth about 300 times. This effectively makes ground-based interferometer arms to be 1200km long (since the variation time of the GW is much longer than the travel time of the laser in the cavity). Accordingly, LIGO is sensitive to frequencies of $f_{\mathrm{LIGO}} \sim 10^2\mathrm{Hz}$. For the future space-based interferometer LISA, the working principle will be the same but with longer arms $L \sim 10^6\mathrm{km}$, being thus sensitive to much smaller frequencies, $f_{\mathrm{LISA}} \sim 10^{-2}\mathrm{Hz}$.

2.2 GWs in cosmology

At large scales, the universe is homogeneous and isotropic to very high accuracy. The background geometry is then described by a (flat) Friedmann-Lemaitre-Robertson-Walker (FLRW) metric

$$ds^2 = g^{\mathrm{B}}_{\mu\nu} dx^\mu dx^\nu = a^2(\eta) \left(-c^2 d\eta^2 + d\vec{x}^2\right)\,, \tag{2.2.1}$$

where $a(\eta)$ is the scale factor and we are timing in conformal time $d\eta = dt/a(t)$. The propagation equation (2.1.14) becomes in Fourier space

$$h''_{ij} + 2\mathcal{H} h'_{ij} + k^2 h_{ij} = 0\,, \tag{2.2.2}$$

where $\mathcal{H} = a'/a$ is the Hubble parameter and primes denote derivatives with respect to conformal time. This is nothing but a wave equation with a friction term due to the cosmic expansion. This Hubble friction will produce a redshift of the frequencies $f^{\mathrm{emit}} = (1+z) f^{\mathrm{obs}}$ and a rescaling of the GW amplitude $h \sim 1/(a \cdot r)$. The previous formulae for a compact binary (2.1.8-2.1.10) written in terms of the observed frequency f^{obs} are thus valid if we replace the chirp mass $\mathcal{M}_c$ by the redshifted chirp mass

$$\mathcal{M}_z = (1+z)\mathcal{M}_c \tag{2.2.3}$$

and the physical distance $a \cdot r$ by the GW luminosity distance

$$d_L^{\mathrm{gw}} = (1+z) \int_0^z \frac{c}{H(z)} dz\,, \tag{2.2.4}$$

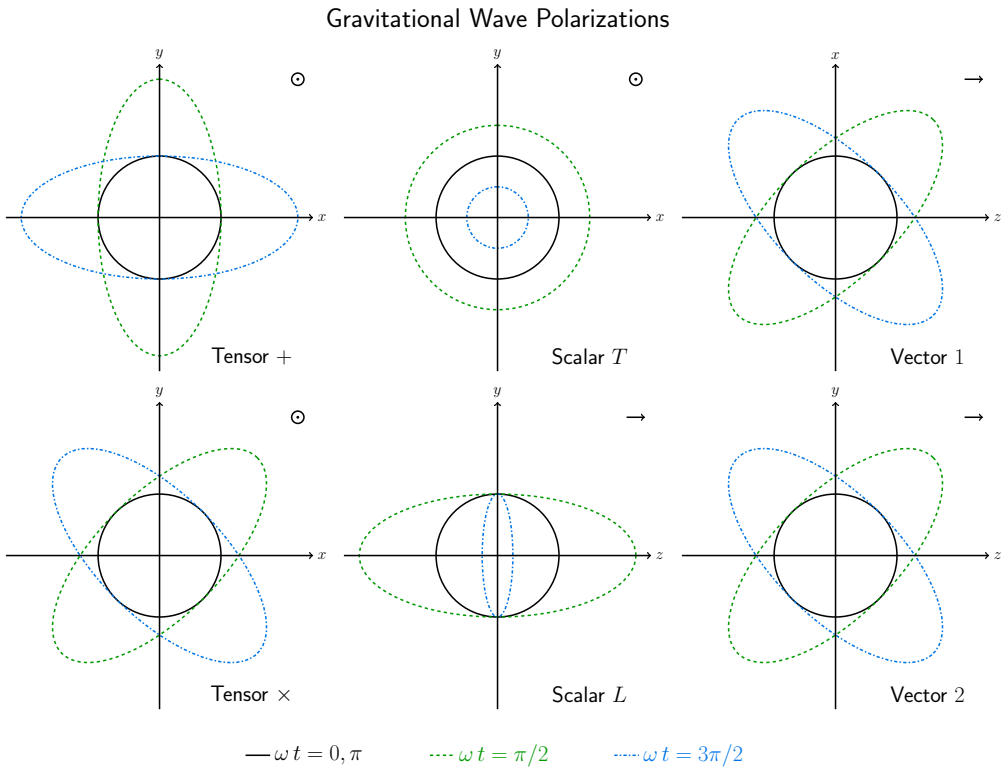

Figure 2.2. Possible gravitational wave polarizations. A circle of test masses is distorted differently for each polarization propagating on the z-direction as a function of time ($\omega t = 0, \pi/2, \pi, 3\pi/2$). General Relativity only contains the two tensor polarizations $+$ and $\times$. Other gravity theories might contain also a transverse (breathing) scalar mode (Scalar T), a longitudinal scalar (Scalar L) and two vector modes (Vector $1\,, 2$).

where c is the speed of light and z the redshift. In this way, all the $(1+z)$ terms cancel each other. Note that there is an intrinsic degeneracy between the redshift and the Hubble parameter $H(z)$ in the GW luminosity distance. Therefore, the expansion history can only be obtained from the GW amplitude if the redshift is known. For near by sources $z \ll 1$, the Hubble constant H_0 can be obtained

$$d_L^{\rm gw} = \frac{cz}{H_0} + \mathcal{O}(z^2)\,, \tag{2.2.5}$$

showing the power of GW astronomy to do cosmology. We will review this topic in more detail in Sec. 2.2.1.

Finally, let us mention that we have only focused on GWs from binary sources in the late universe. However, there could be other sources of GWs in the early universe leading to stochastic, cosmological backgrounds. For a nice review in the subject one can follow [255]. One may wonder if there could be an effect in the GW propagation when traveling through the cold dark matter. This question has been addressed recently and the answer is that the effect is too small [256, 257].

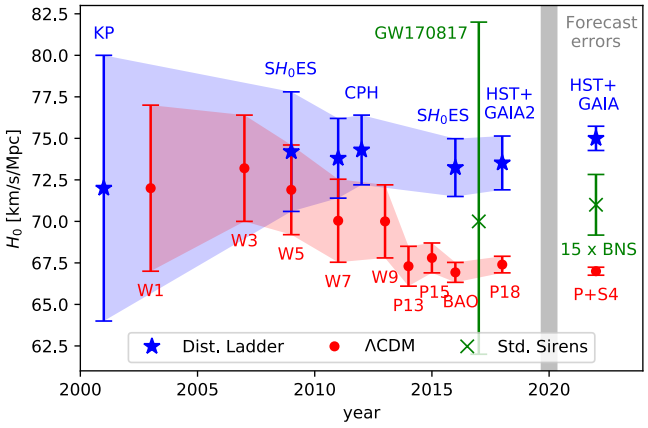

Figure 2.3. The Hubble tension (adapted from [258, 259], including the first standard sirens measurement following GW170817 [33], Planck 2018 [12] and Hubble Space Telescope (HST) with GAIA DR2 [260]). Blue stars correspond to measurements of H_0 in the local universe with calibration based on Cepheids. Red dots refer to derived values of H_0 from the CMB assuming ΛCDM. Green crosses are direct measurements of H_0 with standard sirens. Forecasts are: CMB Stage IV [261], standard sirens [262] and distance ladder with full GAIA and HST [263, 264].

2.2.1 Standard sirens

GWs coming from distant sources can feel the cosmic expansion in the same way as EM radiation does. In fact, we have seen in Sec. 2.2 that the amplitude of the GWs is inversely proportional to the GW luminosity distance $d_L^{\rm gw}$. In GR the GW luminosity distance is equal to EM luminosity distance, with the standard formula given by (2.2.4). However, this is not a universal relation in theories beyond GR as we will discuss in chapter 4. For the moment, we will restrict to Einstein's theory only.

In order to measure distances in cosmology one needs both a time scale and a proper ruler. The inverse dependence of the strain with $d_L^{\rm gw}$ makes GWs natural cosmic rulers. Introducing the full cosmological dependence[2], the GW luminosity distance is given by

$$d_L^{\rm gw} = \frac{(1+z)}{\sqrt{|\Omega_K|}} \text{sinn}\left[c\int_0^z \frac{\sqrt{|\Omega_K|}}{H(z')}dz'\right] , \tag{2.2.6}$$

where $\text{sinn}(x) = \sin(x), x, \sinh(x)$ for a positive, zero and negative spatial curvature respectively. Assuming a ΛCDM cosmology, the Hubble parameter is a function of the matter content Ω_m, the curvature Ω_K and the amount of DE Ω_Λ (radiation at present time is negligible)

$$H(z) = H_0\sqrt{\Omega_m(1+z)^3 + \Omega_K(1+z)^2 + \Omega_\Lambda}\,. \tag{2.2.7}$$

On the contrary, GWs alone do not provide information about the source redshift. This is because gravity cannot distinguish a massive source at large distances with a light

[2] In (2.2.4) we had assumed a flat universe.

source at short distances. Nevertheless, when GWs events are complemented with other signals that allow a redshift identification, these events become *standard sirens* [265]. Standard sirens are complementary to already well-established *standard candles*, SN events in which the intrinsic luminosity can be calibrated allowing for a measurement of the EM luminosity distance. There are also *standard rulers*, such as the one determined by baryon acoustic oscillations (BAO) which provides the angular diameter distance. For binary black-holes (BBH) it is not expected to observe any counterpart, unless there is matter around the BHs [266]. Fortunately, binary neutron stars (BNS) and black-hole neutron star systems (BHNS) are expected to emit short gamma-ray bursts (sGRB) and other EM counterparts, becoming clear standard siren targets.

The first ingredient for a standard siren is the measurement of the GW luminosity distance. However, $d_L^{\rm gw}$ is degenerate with the inclination of the binary. More precisely, showing the explicit angular dependence of the waveform (2.1.8) one finds that the two polarizations scale as

$$h_+ \propto \frac{(1+\cos\iota)^2}{2d_L^{\rm gw}} \quad \text{and} \quad h_\times \propto \frac{\cos\iota}{d_L^{\rm gw}}\,, \tag{2.2.8}$$

where ι is the inclination angle. This distance-inclination degeneracy is the main source of uncertainty of present measurements of $d_L^{\rm gw}$ [267]. One possibility to break this degeneracy is to have an identification of both polarizations. This requires at least a three detector network and a good sky localization. Another possibility to break this degeneracy is when the binary has a precessing spin. Then, there is a characteristic modulation of the amplitude that can disentangle the inclination angle. Orbital precession is more significant for large effective spin $\chi_{\rm eff}$[3] and/or small mass ratios $q = m_2/m_1 \leqslant 1$ since there is also an effective spin-mass ratio degeneracy. Possibly good candidates for this would be BHNS binaries since BNS typically have a mass ratio close to 1.

The other ingredient for a standard siren is the identification of the redshift. This can be achieved by different means:

- *EM counterpart:* the simplest consists in finding an EM counterpart of the GWs from the binary coalescence [265]. Then, the redshift could be extracted from the EM counterpart or from the host galaxy depending on the case. BNS will produce a sGRB after the merger. This sGRB is characterized by a beaming angle θ_j, which is typically expected to be $\theta_j \leqslant 30°$. This means that depending on the orientation of the source we will be able to detect both signals only in a small fraction of the events. Observing a bright afterglow or kilonovae [268] might increase the changes of detecting a counterpart. BNS will be the primary source for LIGO [269], although BHNS could also play an important role [270]. SMBHs might be good standard sirens for LISA as well [271]. Several multi-messenger observations will lead to a precise measurement of the cosmic expansion either for second generation detectors [262, 272] or for third generation [273].

[3]The effective spin is the mass weighted projection of the two spins of the binaries into the orbital angular momentum, see (1.2.2).

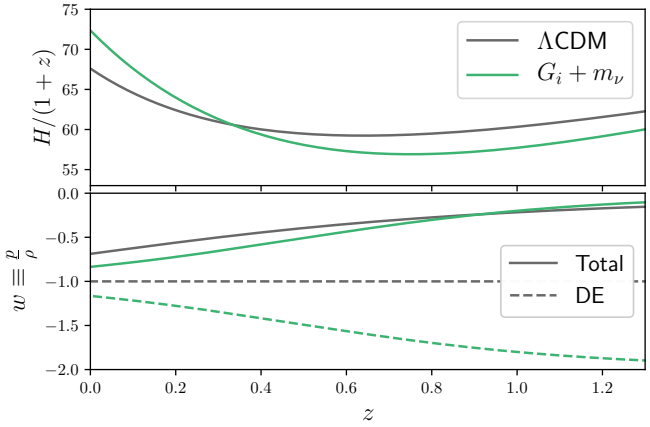

Figure 2.4. Hubble parameter H and equation of state (EoS) w as a function of redshift for the SM of Cosmology (ΛCDM) and for covariant Galileons with massive neutrinos. In the bottom panel, the total EoS $w_{\rm tot} = p_{\rm tot}/\rho_{\rm tot}$ is compared with the EoS of DE $w_{\rm DE} = p_{\rm DE}/\rho_{\rm DE}$.

- *Statistical method:* based on statistical methods, one could associate every GW event with all the galaxies within the error in the localization and compute the cosmology [265, 274]. For a large number of events, the true cosmology will statistically prevail. Conveniently, this method applies to any type of source, including BBH which is the most common observation. Moreover, for very loud (golden) events there might be only few galaxies in the localization box [275]. On the con side, this method relies on a complete galaxy catalogue. Recently this type of *dark siren* analysis was applied to GW170814, a three detector BBH event with good sky localization [34]. The study was applied combining the GW information with the photometric redshift catalog of the Dark Energy Survey.

- *BNS details:* for events involving a NS there are other possibilities. If the EoS of the NS is known, one could compute the tidal effects in the GW phase, which breaks the degeneracy between the source masses and the redshift [276]. A good sensitivity could be achieved with the Einstein Telescope [277]. Since this method relies on the knowledge of the EoS, which most probably will be uncovered through GW observations also, an iterative approach could be performed. In addition, one could benefit from the narrow mass distribution of NS to statistically infer the redshift [278]. Finally, numerical simulations suggests that in BNS a short burst of GWs with a characteristic frequency will be emitted after the merger. If this burst was observed, a redshift measurement could be obtained [279]. The main challenge of this method is possibly the low SNR of the GW burst.

GW170817 has become the first standard siren detected. The redshift, $z = 0.008^{+0.002}_{-0.003}$, was obtained identifying the host galaxy NGC4993 through the different EM counterparts [32]. For such a close event, only the leading term in the cosmic expansion H_0

could be obtained following (2.2.5). The precise value obtained was [33],

$$H_0 = 70.0^{+12.0}_{-8.0}\mathrm{km\,s^{-1}Mpc^{-1}}\,. \tag{2.2.9}$$

This result has the relevance of being the first independent measurement of H_0 using GWs. Still, since it is only one event, the relative error is large, of the order of 14%. From this error budget, 11% arises from the uncertainty in the measurement of the distance due to present detector sensitivity and the previously mentioned degeneracy with the inclination angle. The rest of the error comes from the uncertainty in the estimation of the peculiar velocity of the host galaxy. Observations of the afterglow in different frequencies can help in reducing the inclination uncertainty [280, 281]. One could also use the statistical method to obtain H_0 without information of the counterpart, although the error is significantly larger $H_0 = 76^{+48}_{-23}\mathrm{km\,s^{-1}Mpc^{-1}}$ [282]. Recent studies have shown that with order ~ 50 BNS standard sirens events H_0 could be measured at the level of $\sim 2\%$ [283, 284]. Depending on the actual population of BNS this might be achieved with second generation detectors. LISA will detect mergers of SMBHs (with EM counterparts), providing measurements of cosmic expansion up to $z \sim 8$ and potentially measuring H_0 with $1 - 7\%$ precision [8, 285].

The Hubble rate tension

Standard siren observations of the cosmic expansion can also explore the tension on the Hubble parameter: where a distance ladder measurement gives a value $H_0 = (73.52 \pm 1.62)\mathrm{km\,s^{-1}Mpc^{-1}}$ [260] higher than the model-dependent inference from the CMB $H_0 = (67.4 \pm 0.5)\mathrm{km\,s^{-1}Mpc^{-1}}$ [12] (see in Fig. 2.3). The tension now reaches the level of 3.6σ. Reanalysis of the local distance ladder with more sophisticated statistical techniques tend to agree on the high value, although with somewhat larger error bars [286, 287]. Other low redshift determinations confirm this trend, for instance time delays from multiply-imaged quasar systems [288] give $H_0 = (71.9^{+2.4}_{-3.0})\mathrm{km\,s^{-1}Mpc^{-1}}$. Measurements of H_0 can also be obtained combining BAO and primordial deuterium abundances [289] (see more details in the review [290] and a compilation of the values of H_0 in [291]).

If the tension is not due to systematic errors in either of the surveys, it would indicate a mismatch between the low and high redshift distance ladders [292], which might be the first hint of the need to revise the standard cosmological model. Several partial solutions to the H_0 tension have been proposed, although no satisfactory solution exists. Extensions to ΛCDM have been studied, but no simple model seems to work: for instance, increasing the effective number of relativistic species by $\Delta N_{\mathrm{eff}} \approx 0.4$ eases the tension but enters in conflict with small scale Planck polarization [293], which has been confirmed in the latest Planck results. The role of dark energy (through $w(z)$) has also been investigated in connection with the H_0 tension: no equation of state evolution $w(z)$ can reconcile all datasets, as long as GR holds (although the tension could be eased if BAO or SNe data are not included) [294]. Interacting DE eases the tension, particularly for a phantom-like equation of state with $w \sim -1.2$ [295].

Some dark energy models beyond GR and with massive neutrinos have been proposed to ease the tension. Galileon gravity leads to a phantom-like equation of state (EoS)

$w < -1$ [296, 297]: adding massive neutrinos with total mass $\sum m_\nu \approx 0.6eV$ yielded a good fit to both Planck and the direct H_0 measurement [298]. One should note that although the EoS of Galileons $w_{\rm Gal}$ deviates significantly from $w_\Lambda = -1$, massive neutrinos compensate part of the effect so that the total EoS $w_{\rm tot} = p_{\rm tot}/\rho_{\rm tot}$ is more similar to ΛCDM (see bottom panel of Fig. 2.4). Still, this difference is enough to shift the present value of the Hubble parameter $H_0 \equiv H(z = 0)$ to higher values (see upper panel of Fig. 2.4). A latter analysis, reproduced the result, but found a slight tension with the most recent BAO data [299]. Most importantly, the cosmologically viable Galileons were ruled out by GW speed [5] and weak lensing [300]. Note however that those data employed BAO reconstruction and Galileons are known to affect the non-linear BAO evolution [197], making it more conservative to use the unreconstructed data, for which no tension exists. Non-local gravity has similar features but its less negative equation of state (compensated with $\sum m_\nu \approx 0.3$) leads to a reduced tension rather than close agreement [301].

2.3 GWs beyond GR

As we have emphasized at the beginning of this chapter, the existence of wave solutions for metric perturbations is generic for second order gravity theories. However, the behavior of these GWs can be very different depending on the gravity theory. The differences can arise either at the production or the propagation.

2.3.1 Additional polarizations

During the generation of GWs, the main differences in theories beyond GR is that there could be other polarizations excited. We have seen that in GR only the 2 tensor polarizations propagate (recall (2.1.4)). Nevertheless, modifications of gravity might introduce new degrees of freedom. For instance, in scalar-tensor theories there will be an additional scalar mode. Or in Massive Gravity, where there will be in addition 2 vector modes and a scalar one. For a GW propagating in the z-direction, one could decompose the amplitude A_{ij} in the different polarizations

$$A_{ij} = \begin{pmatrix} A_S + A_+ & A_\times & A_{V1} \\ A_\times & A_S - A_+ & A_{V2} \\ A_{V1} & A_{V2} & A_L \end{pmatrix} , \tag{2.3.1}$$

where A_+ and $A_\times$ are the two tensor modes, $A_{V1,2}$ the two vector polarizations, A_S the transverse (breathing) scalar and A_L the longitudinal scalar mode. One should note that these other types of polarizations will also leave an imprint in the detectors. Each polarization will have a different effect as we exemplify in Fig. 2.2. In principle, with a set of 6 detectors one could distinguish all possible polarizations.

Before continuing, it is important to remark that if a source is emitting additional polarizations, it will lose energy more rapidly. For a binary pulsar, if additional modes were emitted, the orbit would shrink faster due to the higher energy loss. For PSR

B1913+16 (better known as Hulse-Taylor pulsar) [302], the orbit has been tracked for more than four decades now, showing an impressive agreement with GR [303]. Binary pulsars have been intensively used to constrain alternative theories of gravity, placing severe bound on dipolar radiation as reviewed in [304, 305]. An example of this are Einstein-Aether propagating waves [306], which have been constrained from pulsars due to dipolar GW emission [307, 308]. Another would be the constraints on Brans-Dicke from a pulsar-white dwarf binary [309].

Due to these constraints on the emission of additional polarizations, it is usually invoked a screening mechanism around the source to evade them. If this is the case, deviations of GR could only be measured in the propagation of GWs.

2.3.2 Modified propagation

The propagation of GWs in gravity theories beyond GR can be very complicated. The additional fields might modify the background over which GWs propagate and their perturbations could even mix with the metric ones. For simplicity, we will restrict here to cosmological backgrounds. In that case, due to the symmetries of FLRW, tensor perturbations can only mix with other tensor perturbations. Possible deviations from the cosmological wave equation in GR (2.2.2) can be parametrized by

$$h_{ij}'' + (2+\nu)\mathcal{H}h_{ij}' + (c_g^2 k^2 + m^2 a^2)h_{ij} = \Pi_{ij}\,, \tag{2.3.2}$$

where ν is an additional friction term, c_g accounts for an anomalous propagation speed, m is an effective mass and Π_{ij} is a source term originated by the additional fields. For instance, the scalar-tensor analogue of this equation is (1.1.18). It is interesting that the modified GW propagation can also be understood in analogy with optics as GWs propagating in a diagravitational medium [310].

Focusing on the case without sources, $\Pi_{ij} = 0$, the original GR wave-form $h_{\rm GR}$, given by (2.1.8) for instance, will be modified by

$$h_{\rm GW} \sim h_{\rm GR} \underbrace{e^{-\frac{1}{2}\int \nu\mathcal{H}d\eta}}_{\text{Affects amplitude}} \underbrace{e^{ik\int(\alpha_T + a^2m^2/k^2)^{1/2}d\eta}}_{\text{Affects phase}}, \tag{2.3.3}$$

where we have introduced $\alpha_{_T} = c_g^2 - 1$. Mainly, the additional friction will modify the amplitude, while the anomalous speed and the effective mass change the phase. We will discuss the origin and tests of possible modifications of the GW speed and amplitude in chapters 3 and 4 respectively.

For GWs propagating in FLRW backgrounds, a source is present $\Pi_{ij} \neq 0$ when there are additional tensor modes propagating. A paradigmatic example of this is bigravity, where there are two dynamical metrics. In that case, we have to track the evolution of both metric perturbations [316–318]

$$\begin{pmatrix} h'' \\ t'' \end{pmatrix} + \left[k^2 + m_g^2 \begin{pmatrix} \sin^2\theta & -\sin\theta\cos\theta \\ -\sin\theta\cos\theta & \cos^2\theta \end{pmatrix} \right] \begin{pmatrix} h \\ t \end{pmatrix} = 0\,, \tag{2.3.4}$$

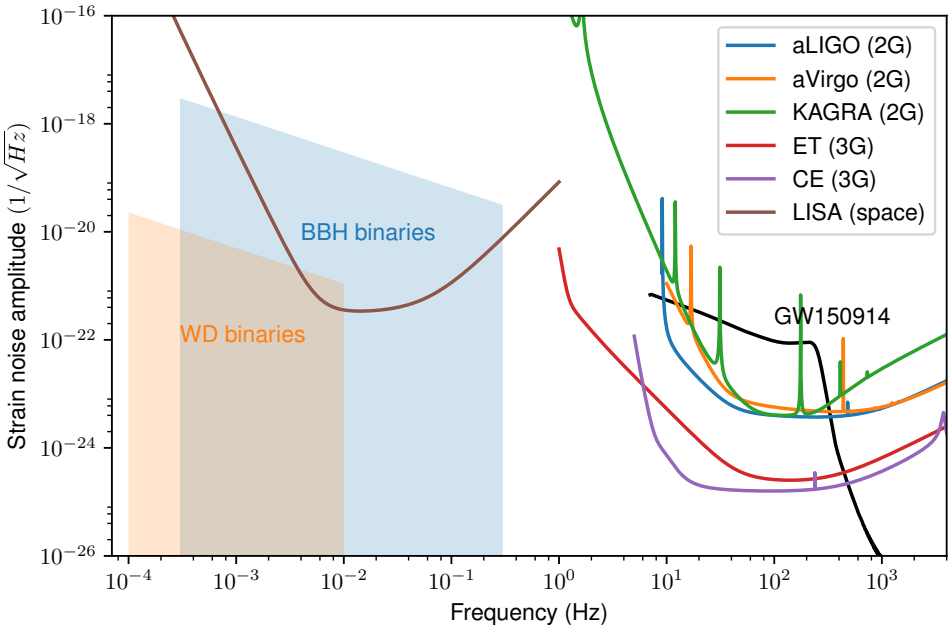

Figure 2.5. Strain sensitivity curves for different GW detectors. Second generation (2G) ground-based detectors are advanced LIGO (aLIGO), advanced Virgo (aVirgo) and KAGRA, with curves given at design sensitivity [311]. Third generation (3G) detectors projected are Einstein Telescope (ET) [312] and Cosmic Explorer (CE) [313]. A space-based detector planned is LISA [314]. For illustration, we include the strain amplitude of GW150914 [254] and the expected background for massive binary black-holes (BBH) and galactic white-dwarf (WD) binaries [315].

where for shortness we have absorbed the Hubble friction in the definition of the perturbation and we do not show the spatial indices. Here m_g is the effective mass (one of the tensor fields is massive) and θ is the mixing angle. Since there are interactions between h_{ij} and t_{ij}, this means that the mass eigenstates are not the same as the propagation eigenstates. In analogy with the propagation of neutrinos, there can be GW oscillations. In chapter 5 we will see how GW oscillations can be tested. One should note that the possibility of having GW oscillations is not restricted to bigravity. Any gravity theory in which the additional degrees of freedom can arrange to form a tensor perturbation over FLRW background could display the same phenomenology. In particular, this is what happens with gauge fields in a SU(2) group [10, 112, 319].

2.4 Present and future GW detectors

Before presenting the different tests of gravity with multi-messenger GW astronomy, let us outline briefly the status of present and future GW detectors. We summarize the different sensitivities of each detector and the typical sources in Fig. 2.5. For illustration, we plot the strain of the first event GW150914 [254]. The capabilities of multi-messenger GW astronomy depend mainly on two aspects:

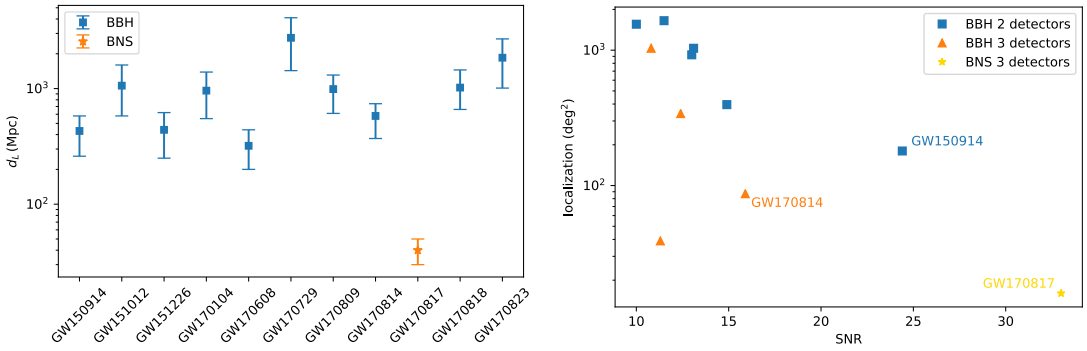

Figure 2.6. On the left, luminosity distance d_L of the GW events reported by the LIGO-Virgo collaboration in runs O1-2 [24]. On the right, 90% credible localization for the 11 events as a function of signal to noise ratio (SNR). We distinguish those events detected only with the two LIGO detectors and those that were found also by Virgo.

- *Number of detections:* this is most sensitive to the size of the volume of the Universe covered by the GW detector. However, there is a large uncertainty in the actual population of the sources, e.g. BNS.

- *Sky localization:* this is most sensitive to the number of detectors that allow for a better triangulation of the source. A better localization of the GW events simplifies the search for a counterpart.

For the GWs events detected so far, we present their luminosity distance and localization in Fig. 2.6. BBHs are observed at the Gpc scale, while the BNS GW170817 was much closer, around 40Mpc. The large uncertainty in d_L is mainly due to the degeneracy with the inclination angle (recall (2.2.8)). Looking to the future, we draft a summary of present expectations for the range of detection and localization angle of different GW detectors in Fig. 2.7. The reader should be aware that these expectations, specially the ones far in the future, might be subject to important modifications.

At present, we are in the second generation (2G) of ground-based detectors. There have been already two operation runs. In the first one, only the two aLIGO detectors were online with a detection range for BNS of the order of 80 Mpc. In the second one, aVirgo joined. Although its sensitivity was still lower, aVirgo helped to reduce the localization area an order of magnitude, from $100-1000\,\mathrm{deg}^2$ to $10-100\,\mathrm{deg}^2$. This can be seen explicitly in the right panel of Fig. 2.6, where we plot the actual localization of the events. The five three-detector events improve significantly the localization w.r.t. the six other two-detector events. The other main factor in the localization is the SNR, to which the localization is inversely proportional to. This exemplifies the expected improvement in the localization with future, more sensitive detectors. In the same manner, loud events seen with several detectors lead to precise localizations, making them perfect targets for statistical identification of the redshift. A good example is GW170814 (see discussion in section 2.2.1).

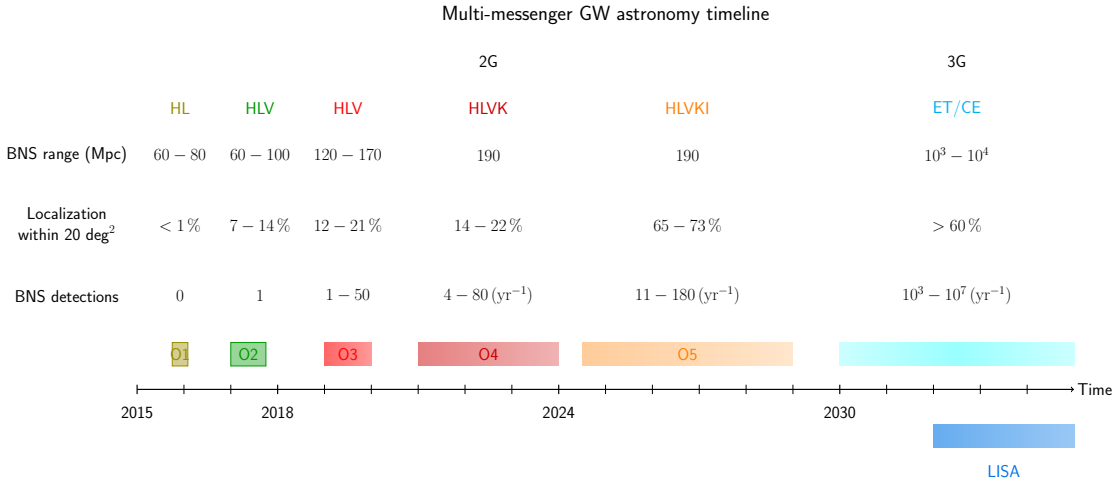

Figure 2.7. Schematic multi-messenger GW astronomy timeline. The binary neutron star (BNS) rate, the localization area in the sky, and the number of BNS detections are given for past and future observation runs. Second generation (2G) ground-based detectors organize in five runs (O1-O5) with different number of detectors online (from 2 to 5) [311]. The nomenclature used is H=Hanford, L=Livingston, V=Virgo, K=KAGRA, I=IndIGO. Third generation (3G) detectors projected are Einstein Telescope (ET) [312] or Cosmic Explorer (CE) [313]. The localization in 3G depends on the network of detectors which is still uncertain [322]. For reference, we include the timeline space-based detector LISA [314]. The reader should note that these numbers correspond to present expectations. For more details we refer to Sec. 2.4.

Neither aLIGO nor aVirgo has reached their designed sensitivity yet. Moreover, other two 2G detectors are on the way. KAGRA [320] in Japan is under construction and it is expected to start operating in 2020. On the other hand IndIGO [321], a replica of LIGO located in India has been approved. This means that in the coming years two main improvements are expected: a larger event rate and a more precise localization [311]. The range of detection is expected to improve by a factor of 3 implying a factor 27 in the detection rate. The localization is expected to reduce to areas of $5 - 20\,\mathrm{deg}^2$ with KAGRA and to a few deg^2 with IndIGO. Note that this is a key point in order to efficiently search for a counterpart or to statistically infer the redshift of the GW event.

A third generation (3G) of ground-based detectors is being planned. The European 3G proposal is the Einstein telescope (ET) [312], an underground, three 10km-arms detector. Its current design aims at improving by a factor of 10 present sensitivity. The US 3G proposal, Cosmic Explorer (CE) [313], is more ambitious with two 40km arms further improving the sensitivity of ET. In any case, 3G detectors imply a substantial change in GW astronomy. While 2G detectors will only be able to reach up to $z \sim 0.05$ for BNS and $z \sim 0.5$ for BBHs, 3G detectors might reach $z \sim 2$ for BNS and $z \sim 15$ for BBHs. In terms of multi-messenger events, this may corresponds to thousands or tens of thousands standard sirens if the EM counterpart is identified. Otherwise, one could use dark sirens.

The sky localization of events in 3G will vary depending on the available network

of detectors [322]. In this sense, there are already plans to upgrade advanced LIGO detectors. This envisioned upgrade is known as LIGO Voyager [323]. Voyager could reach sensitivities between 2G and 3G. The localization will thus vary depending on the redshift of the source since the sensitivity of the network will not be homogeneous. A network of three Voyager detectors plus ET would localize 20% of the events within $10\,\mathrm{deg}^2$, while a setup with three ET detectors would localize 60% of the events within $10\,\mathrm{deg}^2$ [322].

Moreover, space-based GW detectors have been also projected. The European space agency has approved LISA [324]. Being in space and with million kilometer arms, the frequency band and targets of LISA are very different from ground-based detectors (see Fig. 2.5). Expected sources include supermassive BHs, extreme mass ratio inspirals (EMRI) and some already identified white dwarf binaries (known as verification binaries). It is presumed that these sources could be observed with counterparts, enlarging the reach of multi-messenger GW astronomy. For reference, we have included in Fig. 2.5 the expected background of massive BBH ($M_{\rm BH} \sim 10^{4-7} M_\odot$) and unresolved galactic white-dwarf binaries [315] (see more details about the different sources in Fig. 1 of [324]).

Finally, there are other proposals to detect GWs at even lower frequencies, in the band of 1-100 nHz. Sources in this regime could be binary SMBH in early inspiral or stochastic, cosmological backgrounds. These GWs could be observed using a network of millisecond pulsars, in which the pulsation is extremely well-known, for instance with PPTA [325]. Other proposals are to use astrometry with GAIA, which is capable of tracking the motion of a billion stars [326], or to use radio galaxy surveys [327].

2.5 Constraints on modified gravity

In part II of the thesis we are going to explore how to test the propagation of gravitational waves and learn about the underlying gravity theory. In particular, we will focus on the speed of GWs, the GW luminosity distance and the possible GW oscillations in chapters 3, 4 and 5 respectively. In this section we review other tests of gravity that can be performed with GWs that we have not studied during the thesis. Namely, we present how to constraint the mass of the graviton and Lorentz violations with the phase of the GWs. We also show how to test the equivalence principle with multi-messenger observations. Finally, we discuss bounds on additional polarizations.

Mass term

A graviton mass, either effective or fundamental, modifies the propagation speed of GWs. However, contrary to the anomalous speed term c_g, it does it in a frequency dependent way. This means that it can be constrained with GW observation alone, analyzing how the phase of the wave changes in time. The present bound from the LIGO-Virgo collaboration is [328]

$$m_g \leqslant 7.7 \cdot 10^{-23}\,\mathrm{eV}/c^2\,. \tag{2.5.1}$$

Note that this bound is still far away from the cosmologically "motivated" $m_g \sim H_0 \simeq 10^{-33}\mathrm{eV}/c^2$.

Since a graviton mass would also change gravity in other regimes, we can compare the GW bound with other tests. In particular, a massive graviton introduce a Yukawa potential that can be constrained with Solar System observations. This issue has been recently revisited [329], showing that the best bound comes from the perihelion advance of Mars, leading to $m_g < (4-8)\cdot 10^{-24}\mathrm{eV}/c^2$, which is an order of magnitude better than present GW constraints.

For LISA, the GW bound could improve significantly, due to the lower frequencies and higher distances, possibly reaching $m_g < 10^{-26}\mathrm{eV}/c^2$ [330]. In addition, there are proposals to bound m_g measuring the phase lag of continuos sources of GWs and EM radiation with LISA binaries [331–333].[4] For more details in other types of constraints, we recommend the recent review [130].

Modified dispersion relation

Similarly to a graviton mass, Lorentz violating terms modify the dispersion relation in a frequency dependent way. Different wavelengths thus travel at different speeds, modifying the time evolution of GW phase. The effects of the new terms $\mathbb{A}_i$ in the dispersion relation (3.0.2) can be systematically parametrized in modifications of the waveform [334]. A typical example of a Lorentz-violating theory would be high-energy Hořava gravity [46] in which

$$\omega^2 = c^2k^2 + \frac{\kappa_h^4\mu_h^2}{16}k^4 + \cdots\,, \tag{2.5.2}$$

where κ_h and μ_h are parameters of the theory [335].

From the first two events, GW150914 [336] and GW151226 [337], one can already constraints several theories as detailed in Ref. [338]. For Hořava gravity, one can constrain the combination of parameters $\kappa_h^4\mu_h^2$, which were not bounded previously. GW170104 [328] and GW170817 [31] have also been used by LVC to constrain the different $\mathbb{A}_n$.

Equivalence principle

The fact that GWs and EM radiation from GW170817 arrived almost simultaneously at Earth after approximately 100 million light years of travel tells us that both signals follow very similar geodesics. This statement can be made precise in terms of the Shapiro delay [339]. The Shapiro delay measures the difference on arrival time of a massless particle in flat and curved space-time. This can be computed parametrizing the integral of the gravitational potential $U(\mathbf{r})$ over the line of sight [340]

$$\Delta t_S = -\frac{(1+\gamma)}{c^2}\int_{\mathbf{r}_e}^{\mathbf{r}_o} U(\mathbf{r}(l))dl\,, \tag{2.5.3}$$

where $\mathbf{r}_e$ and $\mathbf{r}_o$ are the positions at emission and observation. The prediction of GR is that $\gamma = 1$ for any massless particle. This has been tested to very good precision for photons, $\gamma_{\mathrm{em}} - 1 \leqslant (2.1 \pm 2.3)\cdot 10^{-5}$, using the Cassini space-craft [341]. This is

[4] In fact, one can use the phase lag test to constraint the propagation speed of GWs in general [2].

one of the most stringent Solar System test of gravity and implies that in these scales the gravitational potential should be very similar to GR as discussed in detail in the review [17].

Now, the multi-messenger observation of GW170817 allow us to test if GWs and EM radiation feel the same gravitational potential. In other words, this is testing the equivalence principle. In order to get a bound on the relative difference of $\gamma_{\rm gw}$ and $\gamma_{\rm em}$ one needs to know the gravitational potential between the BNS and the detectors. A conservative bound can be placed introducing only the effect of the Milky Way to arrive at [31]

$$-2.7 \cdot 10^{-7} \leqslant \gamma_{\rm gw} - \gamma_{\rm em} \leqslant 1.2 \cdot 10^{-6} \,. \tag{2.5.4}$$

This constraint has implications for instance for theories in which the dark matter arises from a non-minimal matter coupling to gravity, the so-called dark matter emulators [342]. If both types of waves propagate in the same effective metric, no relative difference is present, as it has been argued for the case in MOG gravity [343].

Additional polarizations

Apart from the modified GW propagation, the other main GW effect of theories beyond GR would be the emission of additional polarizations. We have seen that observing the orbits of pulsars already severely constrains the gravitational energy loss to that of GR. Now, GW astronomy enables to directly probe these extra modes. For this test, the basic role of multi-messenger events is improving the localization and breaking degeneracies with the orientation. [5]

With direct GW observations, the emission of additional polarizations can be constrained from the modifications of the waveforms. For instance, with the first two events it was possible to limit the presence of scalar hair [338]. However, there are still degeneracies between the modified GW phase and the spin and mass parameters that weaken the constraints. This is the case of Einstein-dilaton-Gauss-Bonnet [344] and dynamical Chern-Simons gravity [345], archetypical examples of theories studied in numerical relativity [346, 347].

Moreover, there are also searches for direct signals of non-tensorial polarizations, analyzing the GW geometry through the projection of the different polarizations A_{ij} (2.3.1) onto the detector's network. Since the two LIGO interferometers Hanford and Livingston are basically coaligned, they maximize the SNR of the detection but are insensitive to polarizations. This situation changes with the incorporation of Virgo. From the observation of GW170814, a three detector BBH signal, pure tensor polarization were favored over pure vector or pure scalar modes [348, 349]. However, this was just a simplified analysis and the LIGO-Virgo collaboration is performing a more intensive study including mixed-polarization, which would be a more realistic setup. In the future, these constraints will improve with the switch on of the Japanese detector KAGRA and aLIGO India (see Fig. 2.7). Nevertheless, one should note that quadrupolar detectors

[5] In some sense, one could argue that a simultaneous detection of GR and non-GR polarizations is a multi-messenger observation itself.

like aLIGO and aVirgo cannot distinguish between the breathing and longitudinal scalar modes (see Fig. 2.2).

In addition, it will be possible to test additional polarizations with continuous GW sources such as pulsars [350]. No signal has so far been detected [351, 352], although only the first run has been analyzed because of the costly computational analysis.

Finally, observing the stochastic backgrounds of GWs can probe as well non-GR polarizations. Such background is composed of individually unresolved sources. Since the signal is received from different points in the sky in a continuous manner, it allows a direct measurement of the polarization from the spectral shape of the background [353]. No stochastic GW background has been detected yet, placing limits on the stochastic background from tensor, vector and scalar polarizations [354].

Part II

Gravitational wave tests of Dark Energy

CHAPTER

3

The speed of GWs

The speed of GWs is a fundamental property of any gravity theory. GR predicts that GWs propagate at the speed of light, i.e. $c_g = c$. However, alternative theories may change this prediction. This is because, in contrast to GR, GWs in modified gravity do not have to travel on null geodesics of the background metric (recall (2.1.16)). Generically, one can parametrize at leading order the generalized propagation by

$$\mathcal{G}^{\mu\nu}k_\mu k_\nu + m_g^2 + \sum_{i=4}^{n} A^{\alpha_1\cdots\alpha_n} k_{\alpha_1}\cdots k_{\alpha_n} = 0\,. \tag{3.0.1}$$

Here, $\mathcal{G}^{\mu\nu}$ is the effective metric over which GWs propagate, m_g is the effective mass of the graviton and the tensors $A^{\alpha_1\cdots\alpha_n}$ encode higher-order, wave-vector corrections. When time and space can be decomposed, the above expression leads to a generalized dispersion relation

$$\omega^2 = c_g^2 k^2 + m_g^2 + \sum_{n=4} \mathbb{A}_n k^n\,, \tag{3.0.2}$$

where k is the spatial modulus of the wave-vector and $\mathbb{A}_n$ are the coefficients of the higher order corrections. Accordingly, we can see that the effective metric determines the propagation speed c_g [2] while the higher order wave-vector corrections control Lorentz-violating modifications of the dispersion relation [334]. The mass term m_g also modifies the dispersion relation [355]. Importantly, the speed of GWs c_g leaves a frequency independent effect on the phase of the GW while the mass term and higher-order corrections induce frequency-dependent effects. This implies that the latter can be probed with GW signals alone while the former requires multi-messenger events. Here, we focus on c_g, its origin and its recent multi-messenger constraint.

In this chapter, we will begin in Sec. 5.2.3 showing how to compute the effective GW's metric in generic theories of gravity and discussing under which conditions an anomalous speed arises, i.e. $c_g \neq c$. Then, we will consider two situations in which an anomalous speed is induced: either by the cosmological evolution, Sec. 3.2, or by the spatial variation, Sec. 3.3. We will discuss the impact of these results for dark energy models and theories exhibiting scalar hair or screening mechanisms. Finally, in Sec. 3.4, we will present the implications of the constraints on c_g after the first multi-messenger GW event GW170817: the dead ends (Sec. 3.4.1) and the road ahead (Sec. 3.4.2).

3.1 Conditions for an anomalous speed

The fundamental quantity to determine the causal structure of a gravity theory is the effective metric of each perturbation. The speed of GWs can be determined from the effective metric $\mathcal{G}^{\mu\nu}$ for the tensor perturbations. In theories beyond GR, typically, there will be additional degrees of freedom interacting with the tensor modes. In the following, we will focus on scalar-tensor gravity as a proxy for modifications of gravity, although our arguments could be generalized to other types of fields. We will begin with a formal derivation of the effective metric and then particularize to given theories. Within this class of theories, the second order action for the perturbed fields, $g \to g + h$ and $\phi \to \phi + \varphi$, is given by

$$\mathcal{L} \sim h_{\mu\nu}\, \mathcal{D}_{hh}^{\mu\nu,\alpha\beta}\, h_{\alpha\beta} + h_{\mu\nu} \mathcal{D}_{h\varphi}^{\mu\nu} \varphi + \varphi\, \mathcal{D}_{\varphi\varphi}\, \varphi\,, \tag{3.1.1}$$

where the differential operators $\mathcal{D}$ can be up to second order in derivatives and are built with background quantities.

The first thing we want to achieve is to decouple the leading interactions between the tensor and scalar perturbations. At leading order in derivatives, the tensor equations of motion (EoM) can be written as

$$\mathcal{E}_{\mu\nu} = \mathcal{A}_{\mu\nu}{}^{\rho\sigma\alpha\beta} \nabla_\alpha \nabla_\beta h_{\rho\sigma} + \mathcal{B}_{\mu\nu}{}^{\alpha\beta} \nabla_\alpha \nabla_\beta \varphi + \mathcal{O}\left(\nabla\right)\,, \tag{3.1.2}$$

where, again, $\mathcal{A}_{\mu\nu}{}^{\rho\sigma\alpha\beta}$ and $\mathcal{B}_{\mu\nu}{}^{\alpha\beta}$ are generic tensor of the background quantities. Therefore, if we apply a change of variables

$$h_{\mu\nu} \to h_{\mu\nu} + t_{\mu\nu}\varphi\,, \tag{3.1.3}$$

with $t_{\mu\nu}$ a background tensor satisfying the algebraic equation

$$\mathcal{A}_{\mu\nu}{}^{\rho\sigma\alpha\beta} t_{\rho\sigma} + \mathcal{B}_{\mu\nu}{}^{\alpha\beta} = 0\,, \tag{3.1.4}$$

we can diagonalize the tensor perturbation. This is because we are working at leading order in derivatives (equivalent to leading order in the eikonal approximation [249]) and we are not interested in the additional sub-leading derivative terms that we collect generically in $\mathcal{O}\left(\nabla\right)$. An alternative way to frame this is to say that, since we are interested in local propagation, we adopt Riemann normal coordinates around a point P and expand the scalar and metric background in a Taylor series about P,

$$\begin{aligned} g_{\mu\nu} &= \eta_{\mu\nu} - \frac{1}{3} R_{\mu\rho\nu\sigma} x^\rho x^\sigma + \cdots\,, \\ \phi &= \phi_0 + \phi_\mu x^\mu + \frac{1}{2}\phi_{\mu\nu} x^\mu x^\nu + \cdots\,, \end{aligned} \tag{3.1.5}$$

where $\phi_\mu = \nabla_\mu \phi$, $\phi_{\mu\nu} = \nabla_\mu \nabla_\nu \phi$ and the derivatives and curvatures are all evaluated at P. This leaves freedom for a rotation and boost around P. We may now zoom in and obtain an effective action valid around the point P by taking the scaling limit, $\lambda \to 0$, with

$$x^\mu \to \lambda x^\mu\,, \quad \varphi \to \frac{1}{\lambda}\varphi\,, \quad h_{\mu\nu} \to \frac{1}{\lambda} h_{\mu\nu}\,. \tag{3.1.6}$$

The result is a flat space action, depending on the background field values and derivatives evaluated at P. Therefore, locally, background fields and their derivatives act as constant and do not contribute to the leading derivative part of the perturbations.

As a consequence of the diagonalization, the speed of GWs will be given solely by the action

$$\mathcal{L} \sim h^{\mu\nu}\mathcal{A}_{\mu\nu}{}^{\rho\sigma\alpha\beta}\nabla_\alpha\nabla_\beta h_{\rho\sigma}\,. \tag{3.1.7}$$

In practice, this implies that we only need to focus on the leading derivative terms of the metric in the tensor EoM in order to obtain the effective metric in which the perturbations propagate. However, one has to remember that in the new variables the original tensor and scalar perturbations are mixed.

The next step is to determine the effective metric $\mathcal{G}^{\mu\nu}$ in which the tensor perturbations propagate. For that, we need to project the tensor $\mathcal{A}$ into the second derivatives. The specific form of the effective metric depends on the particular theory. As representative examples, we will study the case of covariant Galileons and scalar Gauss-Bonnet, theories aiming at explaining the late time acceleration and producing scalar hair in black-holes respectively. We will see that in both cases an anomalous speed can appear do to the background configuration of the scalar field.

For a generic scalar and metric backgrounds, the effective medium for the GWs will be *birefrengent*. In others words the effective metric, and thus the speed, will depend on time, the propagation direction and the polarization. We will work in the transverse and traceless gauge of the redefined field, so that

$$\nabla^\mu h_{\mu\nu} = 0 \quad \text{and} \quad h^\mu{}_\mu = 0\,. \tag{3.1.8}$$

After making a temporal slicing, this implies

$$h_{00} = 0 \quad h_{0i} = 0, \quad h_{ij} = h_{ij}^{TT}\,. \tag{3.1.9}$$

The TT-perturbations can be decomposed further in the plus and cross polarizations introducing the polarization tensors $\epsilon_{ij}^{+,\times}$ via

$$h_{ij}^{TT} = h_+\,\epsilon_{ij}^{+} + h_\times\,\epsilon_{ij}^{\times}\,. \tag{3.1.10}$$

Working in this gauge has the usual computational advantages. In particular, it is useful to note how the perturbations of the curvature simplify in the TT-gauge. We will make extensive use of

$$h^{\mu\nu}\delta R_{\mu\nu} \to -\frac{1}{2}h^{\mu\nu}\Box h_{\mu\nu}\,, \tag{3.1.11}$$

$$h^{\mu\nu}\delta R_{\mu\alpha\nu\beta} \to -\frac{1}{2}h^{\mu\nu}\nabla_\alpha\nabla_\beta h_{\mu\nu}\,, \tag{3.1.12}$$

$$\delta R \to 0\,, \tag{3.1.13}$$

when solving the evolution for the linear perturbations.

Since the GW speed will depend on the direction, it is useful to decompose the spatial components of the background tensors in terms of the directions parallel and

perpendicular to the propagation direction of the GW, defined by the wave vector k_i. Specifically, we decompose the spatial gradient as

$$\phi_i = \phi_i^{\parallel} + \phi_i^{\perp} \tag{3.1.14}$$

so that

$$\phi^i \nabla_i = \phi^i_{\parallel} \nabla_i \quad \text{and} \quad \phi^i h_{ij} = \phi^i_{\perp} h_{ij} \,. \tag{3.1.15}$$

In the same manner, we can decompose the spatial part of the second derivatives

$$\phi_{ij} = \phi_{ij}^{\parallel} + \phi_{ij}^{\perp} \tag{3.1.16}$$

so that

$$\phi^{ij} \nabla_i = \phi^{ij}_{\parallel} \nabla_i \quad \text{and} \quad \phi^{ik} h_{kj} = \phi^{ik}_{\perp} h_{kj} \,. \tag{3.1.17}$$

For the transverse component, we can decompose them further

$$\phi^{\perp}_{ij} = \Phi_+ \, \epsilon^+_{ij} + \Phi_\times \, \epsilon^\times_{ij} + \Phi_T \, \epsilon^T_{ij} \,, \tag{3.1.18}$$

where we have introduced the trace part ϵ^T_{ij} of the transverse subspace, i.e. $\phi^{\perp}{}^i_i = \Phi_T$. For convenience, we summarize other relevant projections for the computation using the following notation

$$\phi^2_{\parallel} = g^{ij} \phi^{\parallel}_i \phi^{\parallel}_j = \phi^i_{\parallel} \phi^{\parallel}_i \,, \tag{3.1.19}$$

$$\phi^2_{\perp} = g^{ij} \phi^{\perp}_i \phi^{\perp}_j = \phi^i_{\perp} \phi^{\perp}_i \,. \tag{3.1.20}$$

In particular one could easily show that

$$\phi^i \phi^j \nabla_i \nabla_j = \phi^2_{\parallel} \nabla^i \nabla_i \quad \text{and} \quad \phi^{ij} \nabla_i \nabla_j = \phi_{\parallel}{}^j_j \nabla^i \nabla_i \,. \tag{3.1.21}$$

Moreover, noticing that $h_{ik} h^{kj}$ defines a projector onto the transverse subspace, one can show that

$$h_{ik} \, \phi^k \phi_j \, h^{ij} = \frac{1}{2} \phi^2_{\perp} \, h_{ij} h^{ij} \quad \text{and} \quad h_{ik} \, \phi^k{}_j \, h^{ij} = \frac{1}{2} \phi_{\perp}{}^k_k \, h_{ij} h^{ij} \,. \tag{3.1.22}$$

In addition, let us notice that there could also be full contractions of the tensor perturbations

$$h_{ij} \phi^i \phi^j = h_+ \, \epsilon^+_{ij} \phi^i_{\perp} \phi^j_{\perp} + h_\times \, \epsilon^\times_{ij} \phi^i_{\perp} \phi^j_{\perp} \,, \tag{3.1.23}$$

$$h_{ij} \phi^{ij} = 2(h_+ \Phi_+ + h_\times \Phi_\times) \,. \tag{3.1.24}$$

Importantly, if such terms appear in the EoM, they imply that the propagation would be different for each polarization.

Finally, once we have the effective metric, we can compute the propagation speed c_g. In the case in which the time-spatial components are small, $\mathcal{G}^{0i} \ll \mathcal{G}^{00}, \mathcal{G}^{ii}$, the speed follows analytically from

$$c_g^2 = -\mathcal{G}_{\parallel}{}^i{}_i / \mathcal{G}^0{}_0 \,. \tag{3.1.25}$$

For time-like or space-like gradients, the condition on $\mathcal{G}^{0i}$ is satisfied when the background metric has no mixed timed and spatial components $g_{0i} \ll g_{00}, g_{ii}$ and the scalar background is such that $\phi_{0i} \ll \phi_{00}, \phi_{ii}$.

Conditions for an anomalous speed

We have seen that, at small scales and for arbitrary backgrounds, the action is determined by the effective metric $\mathcal{G}^{\mu\nu}$ over which GWs propagate

$$\mathcal{L} \propto h_{\mu\nu}\mathcal{G}^{\alpha\beta}\partial_\alpha\partial_\beta h^{\mu\nu} = h_{\mu\nu}\left(\mathcal{C}\Box + \mathcal{D}^{\alpha\beta}\partial_\alpha\partial_\beta\right)h^{\mu\nu}\,. \tag{3.1.26}$$

The effective metric can be further decomposed into a piece proportional to the original metric $\mathcal{C}$ and another not proportional $\mathcal{D}$. Then, whenever the (disformal) second term is present, the *GW-cone* will be different from the *light-cone* and both signals will travel at different speeds, as we schematically present in Fig. 3.1. This is because GWs will be traveling on null-geodesics of the effective metric, $\mathcal{G}^{\mu\nu}k_\mu k_\nu = 0$, while light will move on null-geodesics of the background metric $g^{\mu\nu}k_\mu k_\nu = 0$.

In scalar-tensor gravity, two conditions have to be fulfilled to induce an anomalous propagation speed:

i) there is a non-trivial scalar field configuration spontaneously breaking Lorentz invariance. For instance, in order to explain DE, one typically demands $\dot{\phi} \sim H_0$. However, the background profile could also be space-dependent induced by a screening environment or scalar hair.

ii) there is a derivative coupling to the curvature. This highlights the presence of a modified gravity coupling that will lead $\mathcal{D}^{\alpha\beta} \sim \partial^\alpha\phi\,\partial^\beta\phi$.

Whenever these two conditions are satisfied, $c_g \neq c$ and there would be a delay between the GW and the EM counterpart. For instance, differences in the speed of 1%, $c_g/c \sim 0.01$, for sources at 100Mpc induce delays of $\Delta t \sim 10^7$years, clearly beyond human timescales.

Let us examine in more detail the conditions for a disformal relation to arise in a generic theory of gravity. First, it is necessary that the background scalar field has a non-trivial configuration that spontaneously breaks Lorentz invariance. In addition, we note that the effective second-order Lagrangian (3.1.1) follows from the second variation of the action over a background, and is hence equal to the first variation of the EoM. The simplest term in the EoM producing second derivatives and entering in (3.1.26) is the Ricci curvature. When expanded to first-order, considering only the TT components,

$$R^{TT}_{\mu\nu} = -\frac{1}{2}\Box h^{TT}_{\mu\nu} \quad \text{and} \quad R^{TT} = 0\,, \tag{3.1.27}$$

only contribute to the conformal part in the effective gravitational metric (3.1.26).

Further second derivative terms are restricted by covariance to originate either from the Riemann tensor or repeated application of covariant derivatives (e.g. third derivatives of the scalar field), with the two cases related by $\nabla_\mu\nabla_\nu\phi^\alpha = \nabla_\nu\nabla_\mu\phi^\alpha + R^\alpha_{\lambda\mu\nu}\phi^\lambda$. To first-order the TT contribution to the Riemann tensor reads

$$R^{TT}_{\mu\alpha\nu\beta} = -\frac{1}{2}\partial_\beta\partial_\alpha h^{TT}_{\mu\nu} + \frac{1}{2}\partial_\nu\partial_\alpha h^{TT}_{\mu\beta} - (\alpha \leftrightarrow \mu)\,, \tag{3.1.28}$$

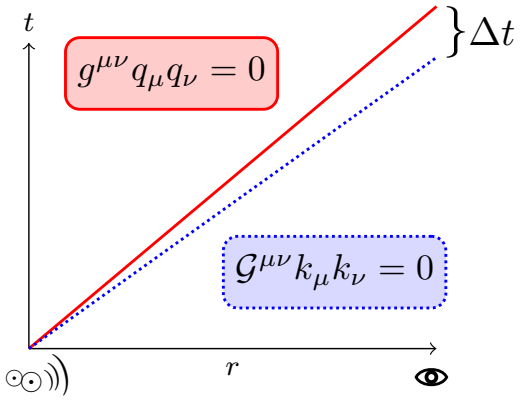

Figure 3.1. Anomalous GW speed. Gravitational waves propagate on an effective metric $\mathcal{G}^{\mu\nu}$ (blue) with a different causal structure than the physical metric $g^{\mu\nu}$ (red) [2]. The speed is derived as $c_g(\vec{k}) = \omega(\vec{k})/|\vec{k}|$ where $k^\mu = (\omega, \vec{k})$ is the solution to $\mathcal{G}^{\mu\nu}k_\mu k_\nu = 0$. Note that the speed can depend on the propagation direction. It may also depend on the frequency (e.g. massive graviton or Lorentz violation), cf. (3.0.2).

The above expression explicitly induces disformal terms in Eq. (3.1.26) via contractions with scalar field derivatives. In quartic Horndeski theories as we will study next in (3.1.31), only ϕ^μ enters in the effective metric (3.1.36) due to the particular non-minimal coupling to the Ricci scalar. In more general cases, for instance when there are couplings to the Ricci tensor such as in quintic Horndeski, second derivatives $\phi^{\mu\nu}$ could appear contracted with the derivatives of the metric and hence in $\mathcal{G}_{\mu\nu}$. Thus, the effective metric would belong to the extended disformal class [3, 96]. In any case, because the Ricci tensor only contributes to the conformal part, the contribution of $R_{\mu\nu\alpha\beta}$ leading to the anomalous speed of GWs is fully captured by the Weyl tensor (i.e. the trace-free part of the Riemann tensor). For a quartic Horndeski theory (3.1.31), the Weyl tensor appears explicitly in the equations of motion whenever $G' \neq 0$ [168].

These considerations allow us to formulate a *Weyl criterion* for anomalous speed of spin-2 GWs. The effective gravitational metric of the example theory (3.1.36) can be generalized to

$$\mathcal{L} \propto h_{\mu\nu}\left(\mathcal{C}\Box + \mathcal{W}^{(\alpha\beta)}\partial_\alpha\partial_\beta\right)h^{\mu\nu}\,, \tag{3.1.29}$$

where $\mathcal{C}$ and $\mathcal{W}^{\mu\nu}$ are the contributions associated with the Ricci and Weyl tensors respectively. Anomalous GW speed requires that $\mathcal{W}^{\alpha\beta} \neq 0$, i.e. for the background scalar derivatives to couple to the Riemann/Weyl curvature. If the Weyl factor is purely time-like and constant around P, $\mathcal{W}^{\mu\nu} = \mathcal{W}^{00}\delta^\mu_0\delta^\nu_0$, the speed of tensors becomes

$$c_g^2 = \frac{\mathcal{C}}{\mathcal{C} - \mathcal{W}^{00}}\,. \tag{3.1.30}$$

In Horndeski theories, which is a general framework that englobes most of the current dark energy models, the EoM are second order [65]. Therefore, the ocurrence of the

Weyl tensor fully distinguishes theories in which $c_g = c$ exactly and those in which the speed of GWs is allowed to vary. GR, Kinetic Gravity Braiding [87] and Jordan-Brans-Dicke theories [72] (including $f(R)$ [79, 80]) only contain Ricci curvature in their equations of motion, and therefore do not modify the speed of GWs. On the other hand covariant Galileons [90] and the covariantization of other generalizations [63, 356–358] will generically predict $c_g \neq c$ [359].

Although the Weyl criterion is characteristic of ST theories, the occurrence of a disformal relation can be applied to more general theories such as massive gravity [137]. In this case the kinetic term has the Einstein-Hilbert form and hence $c_g = c$ plus corrections $\mathcal{O}\left(m^2/E^2\right)$ beyond the scaling limit (3.1.6), as expected from unbroken Lorentz invariance. In the case of bigravity [150] the situation is more subtle, as the kinetic term of the second metric $\sqrt{-f}R[f_{\mu\nu}]$ forces its excitations to propagate along $f_{\mu\nu}dx^\mu dx^\nu = 0$, with $f_{\mu\nu} \neq \Omega(x)g_{\mu\nu}$ in non-flat background space-times. Although matter does not couple to $f_{\mu\nu}$ directly the anomalous speed may be detectable via graviton oscillations, which we will discuss in chapter 5. In vector-tensor theories there could be couplings to the curvature leading to an anomalous propagation speed, for instance $R_{\mu\nu}v^\mu v^\nu$ in vector DE [360]. Interestingly, in more complex vector theories, it is possible to have derivative couplings to the curvature through the field strength $F^{\mu\nu}$ which do not induce an anomalous speed over cosmological backgrounds [112]. This is because in these theories it is possible to have cosmic acceleration while the background of $F^{\mu\nu}$ vanishes, thus violating condition *i*). One should notice that, when violating some of the initial assumptions, the propagation speed of GWs might not be subject to the background value of any additional field and just to the parameters of the theories. This is the case for instance of Hořava gravity [45]. Other theories that attempt to explain away dark matter such as TeVeS also predict an anomalous GW speed [361].

A case of study: quartic Galileons

In the following, we are going to apply our general method to compute the speed of GWs to a quartic shift-symmetric Horndeski theory [65, 85] $S = \int d^4x\ \sqrt{-g}\mathcal{L}$ with

$$\mathcal{L} = G(X)R + G'(X)\left((\Box\phi)^2 - \nabla_\mu\nabla_\nu\phi\nabla^\mu\nabla^\nu\phi\right), \tag{3.1.31}$$

where $X \equiv -\frac{1}{2}(\partial\phi)^2$ and $G' \equiv \partial G/\partial X$. Expanding around a background solution and imposing the transverse gauge condition $\partial^\mu h_{\mu\nu} = 0$, the scaling-limit action reads

$$\mathcal{L} = \frac{1}{2}h_{\mu\nu}\left[G\Box + G'\phi^\rho\phi^\sigma\partial_\rho\partial_\sigma\right]h^{\mu\nu} + h_\mu{}^\rho G'\phi^\mu\phi^\nu\Box h_{\nu\rho} + \cdots, \tag{3.1.32}$$

where we omitted terms involving both the trace of the metric and the scalar field. We then perform a standard 3+1 split of $h_{\mu\nu}$ and restrict to the TT part of the spatial metric components h_{ij}. We will further assume that the spatial shear of the background scalar configuration is negligible.[1] This assumption simplifies the analysis, ensuring that h_{ij}^{TT}

[1] The precise condition is $\phi_{ii} - \phi_{jj}, \phi_{ij} \ll G'/G$ for $(i \neq j)$. This is satisfied in a boosted frame with $\phi_i = 0$ whenever ϕ_μ is time-like.

decouple from the other perturbations and allowing us to ignore the terms omitted in Eq. (3.1.32), which describe the scalar polarization and non-dynamical metric elements.

If the field gradient ϕ_μ is *time-like* (as expected for a cosmological contribution) we can rotate the coordinates so that $\phi_\mu = (\dot{\phi}, 0, 0, 0)$, for some constant $\dot{\phi}$. Then, the last term of (3.1.32) does not contribute and

$$\mathcal{L} = \frac{1}{2}\left\{\left[G - G'\dot{\phi}^2\right]\left(\dot{h}_{ij}^{TT}\right)^2 - G\left(\vec{\nabla}h_{ij}^{TT}\right)^2\right\}, \tag{3.1.33}$$

from which we can read off the propagation speed

$$c_g^2 = \frac{1}{1 - \frac{G'}{G}\dot{\phi}^2}\,. \tag{3.1.34}$$

In particular, GR corresponds to $G(X) = const.$ and we recover $c_g = 1$.

In the case of a *space-like* field gradient we can boost our reference frame so that the time component vanishes. Decomposing the gradient in components parallel and perpendicular to the GW propagation, $\phi_i = \phi_i^{\parallel} + \phi_i^{\perp}$ we obtain that the velocity of propagation of GWs depends on the direction as

$$c_g^2 = 1 + \frac{G'|\phi_\parallel|^2}{G + G'|\phi_\perp|^2}\,. \tag{3.1.35}$$

Thus, for this theory, in general the speed is anisotropic, and the same for the + and × GW polarizations.

As we have seen, the origin of the anomalous speed is in the effective gravitational metric not being conformally related to the original metric, i.e. $\mathcal{G}^{\mu\nu} = \Omega(x)g^{\mu\nu}$. The lack of proportionality is found already in the simple example theory (3.1.31), where

$$\mathcal{G}^{\mu\nu} = G(X)g^{\mu\nu} + G'(X)\phi^\mu\phi^\nu\,, \tag{3.1.36}$$

and $\mathcal{G}^{\mu\nu}$, $g^{\mu\nu}$ are connected by a *disformal relation* [97] for which $\mathcal{G}^{\mu\nu} \neq \Omega(x)g^{\mu\nu}$. Such a relation is ubiquitous in modern scalar-tensor theories [3, 95, 96, 101, 171].

3.2 Cosmological speed and dark energy

In the previous section we have shown how to compute the speed of GWs for generic gravity theories. We found that a non-trivial background field configuration could lead to an anomalous GW speed in certain modified gravity theories. Now, we are going to quantify the difference in the speed in cosmological scenarios. We will find that many models attempting to explain the late time cosmic acceleration indeed have an anomalous propagation speed $c_g \neq c$.

Galileon gravity is a particularly interesting example of a dark energy model that can be thoroughly tested by GW observations. It arises from a scalar field with non-linear derivative self-interactions satisfying the Galilean symmetry $\phi \to \phi + C + b_\mu x^\mu$ in flat space-time [44]. Its covariant generalization [85, 90] is a simple instance of Horndeski's

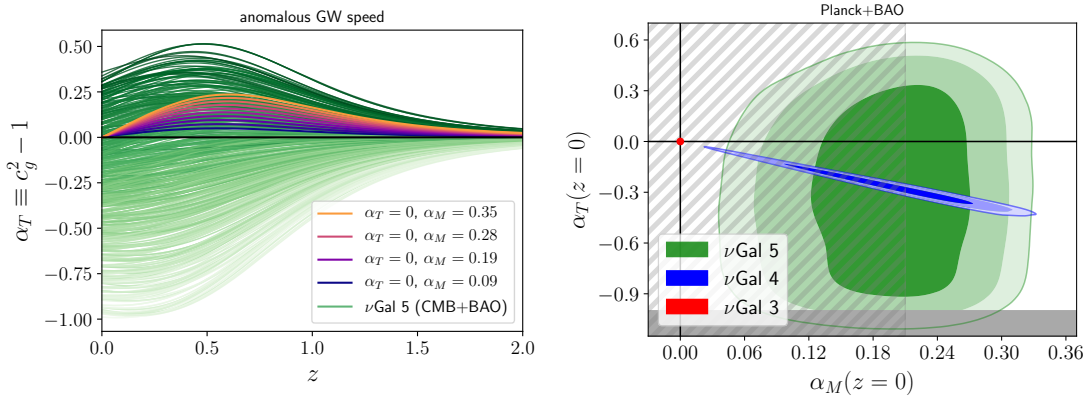

Figure 3.2. **Left:** time evolution of the tensor speed excess α_T as a function of redshift for 300 different realizations of viable quintic Galileon cosmologies. Only quintic fine tuned cases (colored) predict $\alpha_T(z=0) \approx 0$. **Right:** 1, 2 and 3σ confidence regions of the parameter space w.r.t. Planck+BAO for cubic (red), quartic (blue) and quintic (green) Galileons, projected on the $\alpha_T(z=0), \alpha_M(z=0)$ plane. Gray diagonal lines indicate the region disfavored by CMB-LSS cross correlation, measuring the ISW effect (see [299] for details). Models with $\alpha_T < -1$ (gray filled region) have unstable tensor modes.

theory [65], whose action was presented in (A.1.23-A.1.26). The covariant Galileon corresponds to

$$\begin{aligned} G_2(X) &= c_2 X\,, & G_3(X) &= 2\frac{c_3}{M^3}X\,, \\ G_4(X) &= \frac{M_p^2}{2} + \frac{c_4}{M^6}X^2\,, & G_5(X) &= \frac{c_5}{M^9}X^2\,, \end{aligned} \tag{3.2.1}$$

so that all the coefficients of the second-derivative terms are proportional to X. The mass scale $M^3 \equiv M_{\rm Pl}H_0^2$ ensures that the c_i coefficients remain dimensionless ($M_{\rm Pl}$ is the Planck mass). We will refer to three models depending on the highest power of ϕ present in the action: *cubic* ($c_4 = c_5 = 0$), *quartic* ($c_5 = 0$) and *quintic* (all terms).

The covariant Galileon is most interesting as a cosmological model where the Galileon field causes the universe to self-accelerate (without the need of a cosmological constant). As a consequence of shift-symmetry $\phi \to \phi + C$, a tracker solution exists where the time evolution of the field and the Hubble rate obey the relation $\xi \equiv H(t)\dot{\phi}(t)/H_0^2 =$ constant [296]. Under this solution, which has to be reached before DE domination [297], the functions of the modified GW equation (1.1.18) read

$$\alpha_T = \frac{1}{M_*^2 E^4}\left[2c_4\xi^4 + c_5\xi^5\left(1 + \frac{\dot{H}}{H^2}\right)\right]\,, \tag{3.2.2}$$

$$\alpha_M = -4\frac{\dot{H}}{H^2}\frac{M_*^2 - 1}{M_*^2}\,, \quad M_*^2 = 1 - \frac{\xi^4}{E^4}\left(\frac{3}{2}c_4 + c_5\xi\right)\,, \tag{3.2.3}$$

where $E = H(t)/H_0$. Where one should remember that these two functions fully characterizing the GW propagation are: the *tensor speed excess*, α_T, which modifies the

propagation speed of GWs $c_g^2 = 1 + \alpha_{_T}$ and hence the causal structure for this type of signal; and the *running of the effective Planck mass*, $\alpha_{_M} \equiv d\log(M_*^2)/d\log(a)$, which modulates the friction term caused by the universe's expansion.

Self-accelerating Galileon models are all consistent (if massive neutrinos are included) with cosmic microwave background (CMB) and baryon acoustic oscillations (BAO), together with the locally measured value of H_0 (avoiding the tension in ΛCDM) [298,299]. The inclusion of cross-correlations between CMB temperature and galaxies, which probes the Integrated Sachs Wolfe (ISW) effect, trims a significant portion of the parameter space (including all cubic models), but leaves a region that is still viable [299], ($\alpha_{_M}(z = 0) \gtrsim 0.21$). All the cosmologically viable models have an impact of GW propagation [359], as shown in Fig. 3.2.

Stringent bounds can be derived with a multi-messenger GW event constraining c_g. Translated to $\alpha_{_T}$,

$$|\alpha_{_T}| < 9 \cdot 10^{-16} \left(\frac{40\text{Mpc}}{d}\right) \left(\frac{\Delta t}{1.7\text{s}}\right) , \tag{3.2.4}$$

it implies very strong bounds on c_4, c_5. Assuming the non-fine tuned case with no cancellations and noting that $\xi \sim 2$ (range being $1.6 \lesssim \xi \lesssim 3.2$) we find

$$|c_4| \;<\; \frac{\alpha_{_T}}{2\xi^4} \approx 2.8 \cdot 10^{-17} \left(\frac{2}{\xi}\right)^4 , \tag{3.2.5}$$

$$|c_5| \;<\; \frac{\alpha_{_T}}{0.75\xi^5} \approx 3.8 \cdot 10^{-17} \left(\frac{2}{\xi}\right)^5 \tag{3.2.6}$$

(compare with cosmology bounds $c_4 = 0.008^{+0.11}_{-0.026}$, $c_5 = -0.013^{+0.023}_{-0.12}$ at 95% [299]). This in turn constrains the effective Planck mass and its running to be

$$|M_*^2 - 1| < 1.9 \cdot 10^{-15} , \qquad |\alpha_M| < 1.9 \cdot 10^{-15} . \tag{3.2.7}$$

Note that the bounds on M_* and $\alpha_{_M}$ (3.2.7) are specific to Galileon gravity and will in general be independent from those of $\alpha_{_T}$ in other models. With such constraints, the most viable Galileon model in this light is a tiny deviation from the cubic Galileon ($c_4 = c_5 = 0$), which is incompatible with the ISW measurements at 7σ level (note however that generalizations of the cubic Galileon have been shown to fit ISW data [362]). Therefore, combining both GW constraints and cosmological data can rule out covariant Galileons as a DE model. As we will comment later in section 3.4.2, this does not mean that the whole parameter space of Galileons is incompatible, but it does kill the interesting sector of the theory exhibiting self acceleration.

Quintic Galileon models compatible with $\alpha_{_T} \simeq 0$ exist on the very narrow and fine-tuned region of the parameter space where $\Delta t \approx \frac{1}{2}\int_{t_E}^{t_O} \alpha_T(t')dt' \lesssim \Delta t_{obs}$ (Fig. 3.2 left). A second multi-messenger event would, strictly speaking, be necessary to discard this possibility. However, such fine-tuning will not be robust to deviations from the cosmological solution, as we discuss also in section 3.4.2.

3.3 Spatial speed, scalar hair and screening mechanisms

Non-trivial scalar field backgrounds can induce an anomalous GW speed by spontaneously breaking Lorentz invariance. This is not restricted to cosmological setups only. In what follows, we are going to investigate the effect of spatial backgrounds such as scalar hair profiles or screening profiles on the speed of GWs. In both cases, we will limit to backgrounds exhibiting spherical symmetry, meaning

$$ds^2 = -A(r)dt^2 + \frac{dr^2}{B(r)} + r^2 d\Omega^2 \,. \tag{3.3.1}$$

and $\phi = \phi(r)$. Accordingly, for a radial scalar gradient, $\phi_\mu = \phi' \, \delta^r_\mu$, the second derivatives are given by

$$\phi_{\mu\nu} = \phi'' \, \delta^r_\mu \delta^r_\nu - \Gamma^r{}_{\mu\nu} \phi' \,, \tag{3.3.2}$$

where

$$\Gamma^r{}_{\mu\nu} = \text{diag}\left(\frac{1}{2}BA' \,, -\frac{1}{2}\frac{B'}{B} \,, -Br \,, -Br\sin^2\theta\right) \tag{3.3.3}$$

and the $'$ indicates derivatives w.r.t. the radial component. Then, if we take the trace

$$\phi_\mu{}^\mu = B\phi'' + \frac{1}{2}B'\phi' + \frac{1}{2}\frac{A'}{A}B\phi' + 2\frac{B\phi'}{r} \tag{3.3.4}$$

and for the temporal part only $\phi_0{}^0 = +\frac{1}{2}\frac{A'}{A}B\phi'$. Since in general the GW speed will depend on the propagation direction, we will decompose the relevant tensors in the parallel direction $\parallel$ and the perpendicular ones $\perp$. For example, the scalar gradient and the second derivative contractions will vary with the direction:

$$\text{Radial prop.:} \qquad \phi_\parallel^2 = B\phi'^2 \,, \qquad \phi_\parallel{}^i{}_i = B\phi'' + \frac{1}{2}B'\phi' \,. \tag{3.3.5}$$

$$\text{Angular prop. } \theta\text{:} \qquad \phi_\parallel^2 = 0 \,, \qquad \phi_\parallel{}^i{}_i = \frac{B\phi'}{r} \,. \tag{3.3.6}$$

Lastly, we have that the scalar kinetic term is $2X = -B\phi'^2$.

Anomalous GW speed in theories with scalar hair

We begin our analysis by studying how the speed of GWs changes when crossing a region with scalar hair around a compact object, as schematically represented in Fig. 3.3. We parametrize the propagation in terms of an impact parameter b,

$$r = \sqrt{x^2 + y^2} \,, \tag{3.3.7}$$

and an angle

$$\tan\theta = b/x \,. \tag{3.3.8}$$

We will concentrate in the propagation far from the scalar charge Q, at distance much larger than r_h the size of the horizon. In this limit, the background metric will tend to Schwarzschild with

$$A(r \gg r_h) \simeq B(r \gg r_h) \simeq 1 - \frac{2GM}{r} \,. \tag{3.3.9}$$

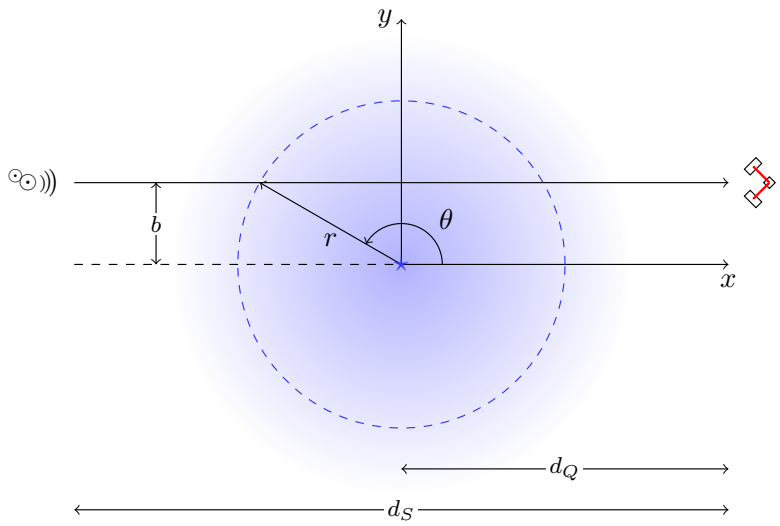

Figure 3.3. GW propagation across a region with a spatial scalar gradient. The source, at a distance d_S from the detector, emits GWs with an impact parameter b w.r.t. the object producing the gradient, which is located at a distance d_Q.

On the other hand, at this scale one is only sensitive to the monopole charge

$$\phi(r \gg r_h) \simeq \frac{Q}{r}\,. \tag{3.3.10}$$

Therefore, since a scalar charge induces a non-trivial spatial profile, there will be an anomalous GW speed in those theories of gravity in which the scalar field couples to the Weyl tensor.

To exemplify how an anomalous speed may arise in theories with scalar hair, we will concentrate on scalar Gauss-Bonnet (sGB) gravity. This theory is the described by the Lagrangian

$$\mathcal{L} = \frac{R}{2} - \frac{1}{2}\nabla_a\phi\nabla^a\phi - V(\phi) + f(\phi)\mathcal{GB}\,, \tag{3.3.11}$$

where $\mathcal{GB} = R^2 - 4R_{ab}R^{ab} + R_{abcd}R^{abcd}$ is the Gauss-Bonnet invariant. This theory can be shown to belong to Horndeski theory choosing[2]

$$\begin{aligned} G_5 &= -4f_\phi \ln X\,, & G_3 &= 4f_{3\phi}X(7 - 3\ln X)\,,\\ G_4 &= 1/2 + 4f_{\phi\phi}X(2 - \ln X)\,, & G_2 &= X + 8f_{4\phi}X^2(3 - \ln X) - V(\phi)\,. \end{aligned} \tag{3.3.12}$$

The tensor EoM are (see for instance [363])

$$G_{ab} = T^{\varphi}_{ab} - 2g_{c(a}g_{b)d}\epsilon^{edjg}\nabla_h\left[{}^*R^{ch}{}_{jg}f'\nabla_e\varphi\right], \tag{3.3.13}$$

[2]See for instance how to obtain this correspondence using the language of differential forms in [1].

where we have used the dual of the Riemann tensor ${}^*R^{ab}_{cd} = \epsilon^{abef}R_{efcd}$. Since we are interested in the leading derivative terms of the tensor perturbations we only have to focus on

$$\begin{aligned}\mathcal{E}^{(2)}_{\mu\nu} =&\frac{1}{2}G_{\mu\nu} + 8(2Xf_{\phi\phi} - [\Phi]f_\phi)R_{\mu\nu}\\ &+8(f_{\phi\phi}\phi^\alpha\phi^\beta + f'\phi^{\alpha\beta})R_{\mu\alpha\nu\beta} - 4(2Xf_{\phi\phi} - [\Phi]f_\phi)Rg_{\mu\nu}\\ &+16(f_{\phi\phi}\phi_{(\mu}\phi^\alpha + f_\phi\phi_{(\mu}{}^\alpha)R_{\nu)\alpha} - 4(f_{\phi\phi}\phi_\mu\phi_\nu + f_\phi\phi_{\mu\nu})R\\ &-8(f_{\phi\phi}\phi^\alpha\phi^\beta + f_\phi\phi^{\alpha\beta})R_{\alpha\beta}g_{\mu\nu}\,.\end{aligned} \tag{3.3.14}$$

Moreover, since we are looking for the TT-perturbations, we can already see that the terms proportional to R or $g_{\mu\nu}$ will not contribute.

For a spatial gradient $\phi_\mu = \phi_i\,\delta^i_\mu$, after applying the appropriate projections described before, we find the wave equation

$$\begin{aligned}&\left(\frac{1}{2} - 8f_{\phi\phi}\phi^2_\parallel - 8f_\phi\phi^0{}_0 - 8f_\phi\phi_\parallel{}^i{}_i\right)\Box h_{\mu\nu}\\ &+ 8\phi^0{}_0 f_\phi(\nabla_0)^2h_{\mu\nu} + 8(f_{\phi\phi}\phi^2_\parallel + f_\phi\phi_\parallel{}^i{}_i)(\nabla_i)^2h_{\mu\nu} = 0\,.\end{aligned} \tag{3.3.15}$$

Then, the GW speed is given by

$$c_g^2 = \frac{1/2 - 8f_\phi\phi^0{}_0}{1/2 - 8(\phi^2_\parallel f_{\phi\phi} + f_\phi\phi_\parallel{}^i{}_i)}\,. \tag{3.3.16}$$

This expression is valid for any space-like scalar gradient and background metric with negligible space-time components mixing. As one can notice, the speed depends on the propagation direction (through $\phi_\parallel$) but is the same for both polarizations.

If we now particularize to a spherically symmetric background and a scalar gradient in the radial direction, the above results simplifies. If the GW propagates in the radial direction, the radial GW speed will be

$$c_r^2 = \frac{1/2 - 4(\ln A)'B\phi' f_\phi}{1/2 - 4B\phi'^2 f_{\phi\phi} - 4(B'\phi' + 2B\phi'')f_\phi}\,. \tag{3.3.17}$$

If the GW propagates in the angular direction θ, the angular GW speed becomes

$$c_\theta^2 = \frac{1/2 - 4(\ln A)'B\phi' f_\phi}{1/2 - 8B\phi' f_\phi/r}\,. \tag{3.3.18}$$

These results agree with Ref. [364], where the radial and angular speed was obtained studying the odd-parity sector of the perturbations of sGB around a BH. This confirms the validity of our method since both approaches are completely independent.[3] Note however that our result (3.3.16) generalize those of [364] allowing for a propagation in any direction.

[3]We have also checked our method by computing the anomalous speed in sGB gravity induced by a time-like gradient. We compare it against the cosmological speed in quintic Horndeski [167], using the appropriate G_i's (3.3.12).

Anomalous GW speed in screened regions

Another situation in which spatial gradients are important is inside a screened region. Specially in theories having a Vainshtein mechanism, since it is based on the spatial derivatives being large [22]. A perfect representative of this type of screening is Galileon gravity, which we have already introduced. Limiting to quartic Galileons, we can use the result of the speed go GWs with a space-like gradient (3.1.35), where we recall that the relevant function

$$G(X) = \frac{M_p^2}{2} + c_4 \frac{X^2}{\Lambda^6} \,. \tag{3.3.19}$$

Here c_4 is a dimensionless, order-1 parameter and Λ the characteristic energy scale of the problem. Then, the speed will be given by

$$c_g^2 = 1 + \frac{2c_4 \, \cos^2\theta \, \phi'^4 \, \Lambda^{-6}}{M_p^2/2 + c_4(1 + 2\sin^2\theta)\phi'^4 \, \Lambda^{-6}} \,, \tag{3.3.20}$$

where we have used the same configuration as before and we have neglected the $2GM/r$ corrections. We observe that the speed varies with the fourth power of the spatial gradient $\phi'^4(r)$. Also, the speed depends on the propagation direction w.r.t. the screening profile. When the GW is perpendicular to the gradient, $\theta = \pi/2$, then the anomalous speed vanishes and one recovers $c_g = 1$. Although in this case the speed is polarization independent, it is interesting to note that for the most general quintic Horndeski c_g will generically depend on the polarization. The fact that the speed in sGB does not depend on the polarization although being a quintic Horndeski model is due to the high degree of symmetry of this theory.

Estimating the spatial time delay

Having all these analytical results in hand, we are now going to estimate the time delay caused by the anomalous GW speed induced by the scalar hair or by the screening background. If we define the zero of the x-axis at the scalar charge, then the time delay between the GW and a EM counterpart will be given by

$$\Delta t = \int_{-(d_S - d_Q)}^{d_Q} \left(c_g^{-1}(x) - c^{-1} \right) dx \,, \tag{3.3.21}$$

where d_S is the distance from the observer to the GW source, and d_Q the distance to the scalar charge Q. This would correspond to the *crossing time delay*. If the source also generates the scalar charge, i.e. $d_S = d_Q$, then the integral would measure the *escape time delay*. It is worth noting that this computation only accounts for the delay induced by one source. In principle, a GW could cross several regions with a non-trivial background. The total delay would then be the sum of the individual ones. In practice a full study would require either knowing the background between the source an the observer or applying statistical methods. We will leave this analysis for future work and concentrate only on the rough estimate of the delay cause by a single source.

Spatial delay induced by scalar hair. In sGB, BHs can develop scalar hair [365, 366]. The charge is determined by the coupling of the theory $f(\phi)$ and the mass of the BH through

$$Q = \frac{4 f_\phi(\phi_\infty)}{M}\,, \tag{3.3.22}$$

where we denote $\phi_\infty = \phi(r \to \infty)$.[4] Using the previous results, at leading order in r for the above background, the inverse GW speed is

$$c_g^{-2} = 1 - \frac{8 f_{\phi\phi} \cos^2\theta Q^2/r^4}{1/2 + 8 f_\phi M Q/r^4}\,. \tag{3.3.23}$$

From this result we can already see that the anomalous speed will be strongly suppressed with the distance, because it is inversely proportional to the fourth power of r. To quantify this departure, we need to choose a particular coupling function $f(\phi)$. We are going to consider two scenarios:

- *Polynomial function:* we parametrize the GB function by a polynomial

$$f(\phi) = \alpha_{\rm GB}\, \phi^n\,, \tag{3.3.24}$$

where $\alpha_{\rm GB}$ is a dimensionful quantity, $[\sqrt{\alpha_{\rm GB}}] = [L]$, parametrizing the strength scale of the sGB coupling. In this way, the derivatives of the coupling function read

$$f_\phi = \alpha_{\rm GB}\, n\, \phi^{n-1} \quad \text{and} \quad \mathrm{f_{\phi\phi}} = \alpha_{\rm GB}\, \mathrm{n(n-1)}\, \phi^{\mathrm{n-2}}\,. \tag{3.3.25}$$

Accordingly, the scalar charge becomes

$$Q = \frac{4\alpha_{\rm GB}\, n\, \phi_\infty^{n-1}}{M}\,. \tag{3.3.26}$$

Altogether, we find that the inverse square of the speed of GWs is

$$c_g^{-2} = 1 - \frac{2^{3(n-1)}\, n^{n+1}(n-1)\phi_\infty^{n(n-1)}\, \cos^2\theta\, \left(\frac{\sqrt{\alpha_{\rm GB}}}{2GM}\right)^{2n+2} \left(\frac{2GM}{r}\right)^{n+2}}{1/2 + 2^{3n-2}\, n^n (n-1)\phi_\infty^{n(n-1)}\, \left(\frac{\sqrt{\alpha_{\rm GB}}}{2GM}\right)^{2n+2} \left(\frac{2GM}{r}\right)^{n+3}}\,, \tag{3.3.27}$$

where for convenience we have normalized the length scales in terms of the BH's Schwarzschild radius $2GM$. From the expression above we see that the anomalous speed tends to c very rapidly with the distance r once $r \gg 2GM$, unless $n < 0$. Therefore, to have an observational effect, we need the BH to be very massive or the coupling $\alpha_{\rm GB}$ to be very large. However, from the observations on the orbital decay of low-mass X-ray binaries [367] and LIGO constraints on wave forms [368], one typically obtains $\sqrt{\alpha_{\rm GB}} \lesssim 10$km.

[4] Note that this relation does not hold for quadratic sGB theory [366] since this theory is defined with $f_\phi(\phi_\infty) = 0$.

- *Exponential function:* we now choose an exponential function

$$f(\phi) = \alpha_{\rm GB}\, e^{\phi}\,, \tag{3.3.28}$$

which is the famous Einstein-dilaton-Gauss-Bonnet (EdGB) theory, inspired originally in string theory. If we take $\phi_\infty = 0$, then the charge is simply

$$Q = \frac{4\alpha_{\rm GB}}{M}\,. \tag{3.3.29}$$

The GW speed is then

$$c_g^{-2} = 1 - \frac{2^9\, e^{\phi}\, \cos^2\theta\, \left(\frac{\sqrt{\alpha_{\rm GB}}}{2GM}\right)^6 \left(\frac{2GM}{r}\right)^4}{1/2 + 2^5\, e^{\phi}\, \left(\frac{\sqrt{\alpha_{\rm GB}}}{2GM}\right)^4 \left(\frac{2GM}{r}\right)^4}\,. \tag{3.3.30}$$

Therefore, we observe the same type of the dependence with the distance: the anomalous speed becomes rapidly suppressed when $r > 2GM$.

Spatial delay induced by screening. For quartic Galileons, the profile of the scalar inside the screened region is given by [22]

$$\phi' \sim \left(\frac{M}{M_p}\right)^{1/3} \Lambda^2 \sim r_V \Lambda^3\,, \tag{3.3.31}$$

where r_V is the Vainshtein radius. Note that for this theory the radial gradient is constant. It is going to be useful to rewrite Λ in terms of a length scale

$$\Lambda \equiv \left(\frac{M_p}{L^2}\right)^{1/3}\,. \tag{3.3.32}$$

In this way, we notice that the GW speed becomes

$$c_g^2 = 1 + \frac{2c_4\, \cos^2\theta\, (r_V/L)^4}{M_p^2/2 + c_4(1 + 2\sin^2\theta)(r_V/L)^4}\,. \tag{3.3.33}$$

Therefore, the anomalous speed will be proportional to

$$\alpha_{\rm gw} = c_g^2 - 1 \sim \left(\frac{r_V}{L}\right)^4\,. \tag{3.3.34}$$

Thus, the degree of deviation w.r.t. the speed of light will depend on the ratio between the Vainshtein radius and the length scale associated to the Galileon interactions.

Let us now make some estimates. In terms of the Schwarzschild radius for a solar mass, the ratio between these two length scales becomes

$$\frac{r_V}{L} \sim \left(\frac{M}{M_\odot}\right)^{1/3} \left(\frac{3{\rm km}}{L}\right)^{1/3}\,. \tag{3.3.35}$$

We could compare this with the usual cosmological Galileon scale in which

$$\frac{r_V}{L} \sim 10^{-7} \left(\frac{M}{M_\odot}\right)^{1/3} \left(\frac{cH_0^{-1}}{L}\right)^{1/3} . \tag{3.3.36}$$

Therefore, the sun has a Vainshtein radius of the order of the kpc. Going back to the speed, we find

$$\alpha_{\rm gw} \sim 10^{-16} \left(\frac{M}{10^9 M_\odot}\right)^{4/3} \left(\frac{cH_0^{-1}}{L}\right)^{4/3} . \tag{3.3.37}$$

As a consequence, in the case of cosmological Galileons, the screening of galaxies can in principle induce deviations in the speed that are measurable. If instead we modify the scale of Galileons, we obtain

$$\alpha_{\rm gw} \sim 10^{-16} \left(\frac{M}{M_\odot}\right)^{4/3} \left(\frac{\rm pc}{L}\right)^{4/3} . \tag{3.3.38}$$

Thus, if L is smaller than a parsec the deviation in $\alpha_{\rm gw}$ might also be detectable with multi-messenger events.

3.4 Constraints after GW170817

Prior to the direct detection of GWs, there were indirect constraints on the speed of GWs. High energy cosmic rays from galactic origin set a stringent lower bound $-2 \cdot 10^{-15} \leqslant c_g/c - 1$ [369], due to the absence of gravitational Cherenkov radiation [370]. The reason is that if gravitons propagate slower than the speed of light, cosmic rays could decay into them and their signal would be lost. This lower bound affects Horndeski theory [371]. However, note that we are talking about very energetic gravitons, different from the low energy GW emission of an astrophysical compact binary. Moreover, the GW speed was indirectly constrained at the level of $|c_g/c - 1| \leqslant 0.01$ with the orbits of binary pulsar in the absence of screening of the cosmological solution [372].

With the detections of GWs from BBHs, the first direct constraints on the speed of GWs were placed [373]. The constraints were still not very strong, $-0.45 \leqslant c_g/c - 1 \leqslant 0.42$, due to the uncertainties in the localization of the source and the low number of detections (3 at the time of the analysis). Detecting a GW with an EM counterpart changes the situation completely, leading to very precise measurements.

Such a multi-messenger GW event was detected on August 17, 2017 with the BNS GW170817 [374]. The GW signal was followed by a short gamma ray burst (sGRB) only $\Delta t = 1.74 \pm 0.05$s after [31]. The source was localized at a distance of $d_{\rm L} = 40^{+8}_{-14}$Mpc. In order to set the constraints, the LIGO-Virgo collaboration conservatively considered the source at the lowest distance of 26Mpc. For the upper bound, it was assumed that both the GW and the sGRB were emitted at the same time and that all the delay is caused by the faster propagation of the GW. For the lower bound, they assumed that the sGRB was generated 10s after the GW, order of magnitude expected in standard astrophysical

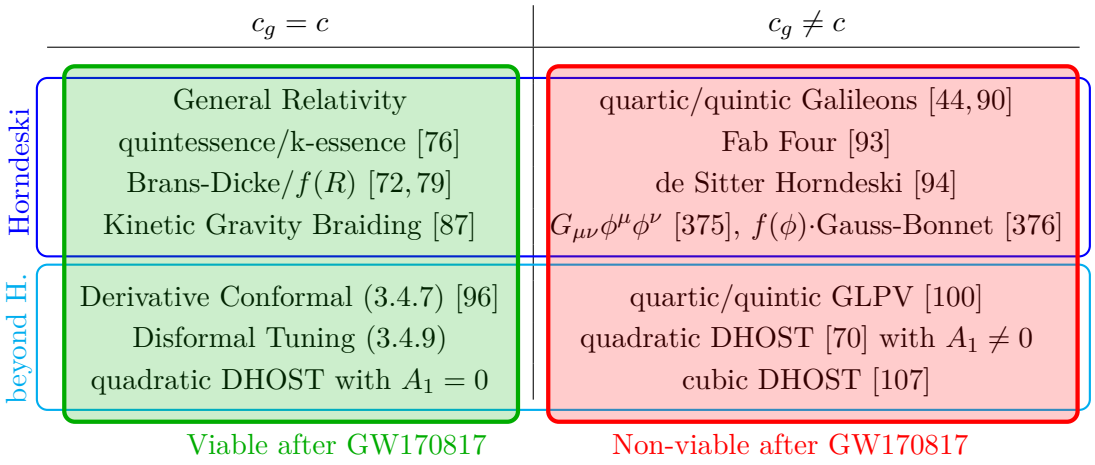

Figure 3.4. Summary of the viable (left) and non-viable (right) scalar-tensor theories after GW170817. Only simple Horndeski theories, $G_{4,X} \approx 0$ and $G_5 \approx$ constant, and specific beyond Horndeski models, conformally related to $c_g = 1$ Horndeski or disformally tuned, remain viable.

models, and that the delay was reduced to 1.74s due to the slower propagation of the GW. In total, this led to the impressive constraint

$$-3 \cdot 10^{-15} \leqslant c_g/c - 1 \leqslant 7 \cdot 10^{-16}\,. \tag{3.4.1}$$

This result has profound implications for many gravity theories and dark energy models.

3.4.1 Dead ends

An anomalous speed arises when there is background field spontaneously breaking Lorentz invariance at the same time that there are derivative couplings to the curvature. If we want a theory to have luminal propagation, at least one of the conditions for an anomalous GW speed has to be broken. If we want the scalar field to keep playing a role in the cosmic expansion history, it cannot have a trivial scalar field configuration. Therefore, the only possibility to satisfy GW170817 is to break the second condition an eliminate derivative couplings to the curvature. For Horndeski theory (A.1.23-A.1.26) this implies [5, 99, 377, 378]

$$G_{4,X} \approx 0\,, \quad G_5 \approx \text{constant}\,. \tag{3.4.2}$$

Translating this result, only the simplest models such as quintessence, Brans-Dicke or Kinetic Gravity Braiding survive. On the contrary, models like Covariant Galileons, Fab Four, Gauss-Bonnet or some sectors of beyond Horndeski are ruled out as DE models. The fact that the parameter space has been drastically reduced has implications for cosmological constraints [184, 213, 379] and for large scale structure [380]. In Fig. 3.4 we present a summary table of the scalar-tensor models that are viable and non-viable after GW170817. Some of the viable models will be discussed in the next section 3.4.2.

For vector-tensor theories the situation is very similar. In order to describe DE and to pass the GW test some couplings of the theory have to be eliminated [5, 377],

in particular $G_{4,Y} \approx 0$ and $G_{5,Y} \approx 0$ (see full action in Eq. (299) of [61]) The same happens for Hořava gravity where one has to impose $\xi \approx 1$ or $\beta_{kh} \approx 0$ [381], which correspond to the conditions for the low-energy version of the theory or its Einstein-aether analogue respectively. The implications of GW170817 for other gravity theories have been extensively explored, for instance for doubly-coupled bigravity [382], $f(T)$ gravity [383] or Born-Infeld models [384]. In Fig. 1.2 we presented a schematic roadmap of the different modifications of gravity highlighting which of them are constrained by the speed of GWs. This figure exemplifies graphically the large impact that GW170817 have had on cosmological modifications of gravity.

3.4.2 The road ahead

We have found that the appearance of an anomalous speed, $c_g \neq c$, can be understood in terms of an effective geometry for the tensor perturbations $\mathcal{G}^{\mu\nu}$, with a different causal structure than the metric field $g^{\mu\nu}$. For Horndeski and beyond Horndeski it has the following form

$$\mathcal{G}^{\mu\nu} = C g^{\mu\nu} + D\phi^{,\mu}\phi^{,\nu} + E\phi^{,\mu\nu}\,, \tag{3.4.3}$$

where the coefficients depend on ϕ and its derivatives, and all quantities are local. GWs propagation is determined by the on-shell *GW-cone* condition $\mathcal{G}^{\mu\nu}k_\mu k_\nu = 0$, for $k_\mu = (\omega, \vec{k})$, and the propagation speed is $c_g^2(\vec{k}) = \omega^2(\vec{k})/k^2$. The anomalous GW speed occurs whenever $\mathcal{G}^{\mu\nu} \neq \Omega(x) g^{\mu\nu}$, i.e., $D, E \neq 0$ in (3.4.3). Quartic theories (A.1.25) produce D-type terms [2], while quintic theories (A.1.26) produce also E-type terms [385]. Both D, E terms can be associated to the presence of the Weyl tensor in the equations of motion [2].

Satisfying the bound $c_g/c - 1 < 10^{-15}$ requires either both operators leading to D, E to be very suppressed, or an internal cancellation between different terms. But such a cancellation is robust against perturbations only if the different terms involved have the same tensor structure, i.e. different terms contributing to D cancel among themselves and likewise for E. In contrast, a cancellation between D and E at the level of the cosmological solution is broken by the presence of perturbations. Assuming that such a cancellation exists, computing the effective metric over a perturbed scalar-field solution $\phi = \bar{\phi}(t) + \varphi(x)$ leads to $\mathcal{G}_{\mu\nu}k^\mu k^\nu = C(\vec{k}^2 - \omega^2) - 2E\omega\vec{k}\cdot\vec{\partial}\dot{\varphi} + \cdots$ after boosting so $\phi_\mu = (\omega, \vec{0})$. The GW speed then depends on the direction and can not be compensated

$$c_g^2 = \frac{\omega^2}{k^2} = 1 + 2\frac{E}{C}\hat{k}\cdot\vec{\partial}\dot{\varphi} + \cdots\,, \tag{3.4.4}$$

where the ellipsis denotes terms that modify the GW speed isotropically. Although this is a second order effect, this deviation will be highly constrained. Thus, tuning the cosmological evolution is not a viable solution to avoid the GWs speed constraint.

After GW170817, finding viable theories amounts to suppressing or compensating the terms leading to a different causal structure, i.e. D and E. In the framework of Horndeski the only option is to suppress the terms leading to an anomalous speed, i.e. (3.4.2). Note that a cancellation of the anomalous speed between G_4 and G_5 will not

be possible in general because they contribute independently to one D and one E term in (3.4.3). The above condition is satisfied only by the simple models contained in $G_2(X,\phi), G_3(X,\phi), G_4(\phi)$.

Viable theories beyond Horndeski can be obtained by modifying the causal structure of the gravitational sector. This can be achieved by applying a disformal transformation of the metric $g_{\mu\nu} \to \tilde{g}_{\mu\nu}$, where

$$\tilde{g}_{\mu\nu} = \Omega^2(\phi, X) g_{\mu\nu} + \mathcal{D}(\phi, X)\phi_{,\mu}\phi_{,\nu}\,, \tag{3.4.5}$$

which changes the GW-cone whenever $\mathcal{D} \neq 0$. Accordingly, the speed of GWs transforms to[5]

$$\tilde{c}_g^2 = \frac{c_g^2(\tilde{X})}{1 + 2\tilde{X}\mathcal{D}}\,, \tag{3.4.6}$$

where c_g is the speed of tensors of the original gravity theory and $-2\tilde{X} = \tilde{g}^{\mu\nu}\phi_{,\mu}\phi_{,\nu}$. This result leaves us with two ways to construct gravity theories with GWs moving at the speed of light:

1) start with a theory with $c_g = 1$ and apply a conformal transformation, $\mathcal{D} = 0$, or

2) compensate the anomalous speed with a disformal factor, i.e. $\mathcal{D} = (c_g^2 - 1)/2\tilde{X}$.

Starting with a $c_g = 1$ Horndeski theory and applying a conformal transformation leads to

$$\mathcal{L}_C = \frac{1}{16\pi G}\left(\Omega^2 R + 6\Omega_{,\alpha}\Omega^{,\alpha}\right) + \tilde{\mathcal{L}}_2 + \tilde{\mathcal{L}}_3\,, \tag{3.4.7}$$

with $\Omega = \Omega(X,\phi)$ and where $\tilde{\mathcal{L}}_i$ are the transformed Horndeski $\mathcal{L}_2, \mathcal{L}_3$ (A.1.23-A.1.24) (which transform into combinations of themselves under a disformal relation (3.4.5)). The above theory (3.4.7), first presented in Ref. [96], was latter identified as a DHOST theory [70] and hence ghost-free. It includes mimetic gravity [98] as a particular case.

Compensating the anomalous speed may also render a theory viable. For a quartic Horndeski theory (A.1.25) with $c_g^2(X) = G_4/(G_4 - 2XG_{4,X})$ (recall (3.1.34)), one needs a beyond Horndeski GLPV Lagrangian [100]

$$\begin{aligned}\mathcal{L}_4^{bH} = F_4(\phi, X)\big(&\phi_{,\mu}\phi^{;\mu\nu}\phi_{;\nu\rho}\phi^{\rho} - \phi_{,\mu}\phi^{\mu\nu}\phi_{,\nu}\Box\phi \\ &- X\left((\Box\phi)^2 - \phi_{;\mu\nu}\phi^{;\mu\nu}\right)\big)\,.\end{aligned} \tag{3.4.8}$$

This term introduces an extra contribution to the speed of gravitational waves that can be used to tune away the anomalous GW speed:

$$c_g^2 = \frac{G_4}{G_4 - 2X(G_{4,X} - XF_4)} = 1 \Leftrightarrow F_4 = G_{4,X}/X\,. \tag{3.4.9}$$

[5] We apply the disformal transformation (3.4.5) to the gravity sector only. A field redefinition of the whole action, including matter, will not change the physical ratio c_g/c. Note that dependence of the transformation coefficients in X will introduce beyond Horndeski terms in the action (A.1.23-A.1.26) [95].

Not surprisingly, the combined theory is the result of applying a disformal transformation (3.4.5), with a suitably chosen $\mathcal{D}$, to the starting Horndeski theory. It is important to emphasize that this particular cancellation holds over general backgrounds, as it involves D-terms in the effective metric (3.4.3). Our results agree with the independent derivation presented in Ref. [99].

Thus, the most general ST theory with $c_g = 1$ is given by

$$\mathcal{L}_{c_g=1} = \mathcal{L}_C + \mathcal{L}_4 + \mathcal{L}_4^{bH} \tag{3.4.10}$$

given by Eqs. (A.1.25), (3.4.7) and (3.4.8)), subject to the compensation condition (3.4.9) (note that the conformal theory contains Horndeski's G_2, G_3, and $G_4(\phi)$). This can be understood in the framework of quadratic DHOST theories [104, 386] for which

$$c_g^2 = \frac{G_4}{G_4 + 2XA_1} \tag{3.4.11}$$

(for a cosmological background with a timelike scalar gradient) where A_1 is the coefficient of the $\phi_{;\mu\nu}\phi^{;\mu\nu}$ terms in the action. It is very easy to see that this term is canceled by the combination such that the compensation (3.4.9) holds. Note that, as in Horndeski and GLPV, terms with higher powers of $\nabla\nabla\phi$, cubic DHOST [107] in this case, cannot help in erasing the anomalous GW speed since they contribute to different terms in the effective metric (3.4.3).

Before concluding, let us make some remarks about the implications of the constraints on the speed of GWs for gravity and DE:

- *Constraints apply to dark energy models:* it is important to emphasize again that the constraints after GW170817 generically apply only to gravity theories in which the additional fields have a relevant role in cosmology. For instance, we could take as an example the case of quartic Galileons. From (3.2.2) we can see that the tensor speed speed excess is proportional to

$$(c_g^2 - 1) \propto \left(\frac{\dot{\phi}}{H_0}\right)^4 . \tag{3.4.12}$$

  Accordingly, for models in which the Galileon triggers the present expansion, $\dot{\phi} \sim H_0$, there are $\mathcal{O}(1)$ deviations in c_g. However, if we resign from this goal, we could be in agreement with GW170817 simply choosing $\dot{\phi} < 10^{-4}H_0$. Of course, this sector of the theory is less interesting a priory but serves to exemplify that GW170817 constrains DE models rather than gravity theories. Another example considered recently is scalar Gauss-Bonnet gravity [387], where it was shown that if the scalar does not have a dominant energy density, it passes the constraints on the speed easily.

- *Cosmological tunning of $c_g(z = 0) = c$ is not viable:* although tunning by hand the parameters of the theory to pass GW170817 is not appealing, it is tempting to devise a dynamical mechanism that leads to this tunning. One could therefore

think that a cosmological model with such mechanism could be viable. However, as we have stressed along the discussion, to avoid the constraint on the speed of GWs, one has to fix $c_g = c$ on arbitrary backgrounds. This is because, on real life, when the GW travels from the source to the observer, it will cross backgrounds deviating from FRW, e.g. when they cross the Milky Way or due to the large-scale structure. Therefore, delays between the GW and the EM radiation will be accumulated again. Given the strength of the GW170817 constraint, any small deviation from the cosmological background would kill the tunning mechanism. This reasoning has been applied lately to an interesting sector of Horndeski gravity in which the scalar EoM dynamically cancel the anomalous speed [388]. Nevertheless, large-scale inhomogeneities are sufficient to make the mechanism fail.

- *Constraints assume EFT validity:* when computing the speed of GWs or any other GW observable from a dark energy model, we are assuming that the effective field theory is valid. In other words, we are assuming that higher order operators does not modify our action. However, one should note that the frequency of GW170817 was close to the typical strong coupling scale of the EFT of DE

$$\Lambda_{\rm strong} \sim (M_{\rm pl} H_0^2)^{1/3} \sim 260\,{\rm Hz}\,. \tag{3.4.13}$$

If the cutoff of the theory is of the order of the strong coupling scale $M_{\rm cutoff} \sim \Lambda_{\rm strong}$, as it is usually assumed, higher dimensional operators might modify the dispersion relation although one would not expect that they conspire to completely cancel the anomalous speed at the level of $\mathcal{O}(10^{-15})$ [99]. In the case in which the cutoff scale is parametrically smaller, $M_{\rm cutoff} \ll \Lambda_{\rm strong}$, the situation could be different [389]. Theories with a Lorentz-invariant ultra-violet (UV) completion are presumed to have luminal GW propagation. Therefore, one would expect higher dimensional operators to erase any anomalous speed beyond the cutoff scale, which in this case might already happen in the LIGO band. However, the speed of GWs cannot be computed beyond $M_{\rm cutoff}$ if the UV completion is unknown. In any case, the hypothesis that higher dimensional operators render $c_g(k_{\rm LIGO}) = c$ could be tested detecting GWs at different frequencies, for example with LISA (see Fig. 2.5). A way to test an anomalous GW speed at LISA frequencies, $c_T(k_{\rm LISA}) \neq c$, is to measure the phase lag between GW and EM signals of continuous sources such as the LISA verification binaries. This test can constrain the graviton mass [331, 332] as well as the propagation speed [2, 333]. In appendix B, we detail how to perform this test. Measuring c_g at a different frequency might give us valuable information about the cutoff scale of the effective theory of DE.

CHAPTER

4

GW luminosity distance

Apart from the speed of GWs, the other main observable from the modified propagation is the luminosity distance of GWs, $d_L^{\rm gw}$. Assuming that the emission of GWs is not modified w.r.t. GR, we can use the scaling of the amplitude of the GW to test the propagation. We can obtain the GW luminosity distance from the inverse of the amplitude (recall Eq. (2.1.8)),

$$|h_{+,\times}| = \frac{\mathcal{M}_z^{5/3} f^{2/3}}{d_L^{\rm gw}} F_{+,\times}\,, \tag{4.0.1}$$

where $\mathcal{M}_z$ is the redshifted chirp mass, $\mathcal{M}_z = (1+z)\mathcal{M}_c$, f the frequency and $F_{+,\times}$ is a polarization dependent function of the inclination angle. We have seen in (2.2.2) that GR predicts that GWs are only sensitive to the Hubble friction due to the cosmic expansion. Accordingly, the GW luminosity distance is equal to the EM one, which is determined by the Hubble parameter

$$d_L^{GR} = d_L^{\rm em} = (1+z)\int_0^z \frac{c}{H(z)} dz\,. \tag{4.0.2}$$

However, in other theories of gravity, the cosmic medium could be, for instance, more absorptive. This would dim the received signal, which would be interpreted as the source being further apart. In other words, the GW luminosity distance $d_L^{\rm gw}$ would be larger than the EM luminosity distance $d_L^{\rm em}$.

A paradigmatic example of a modification of gravity in which the GW luminosity distance differs from the EM one is adding extra dimensions [390]. In extra dimension theories, for instance DGP, there can be a large distance leakage of the gravitons into the additional dimensions. This means that, as a net effect, an observer will receive less gravitons or, in other words, the gravitational signal will be dimmer. By dimensional analysis, the GW luminosity distance scales in these theories as

$$\frac{d_L^{\rm gw}(z)}{d_L^{\rm em}(z)} \propto (d_L^{\rm em}(z))^{(D-4)/2}\,, \tag{4.0.3}$$

where D refers to the number of space-time dimensions in which the graviton can propagate. For $D = 4$, one recovers the GR result $d_L^{\rm gw} = d_L^{\rm em}$. In cases in which the graviton

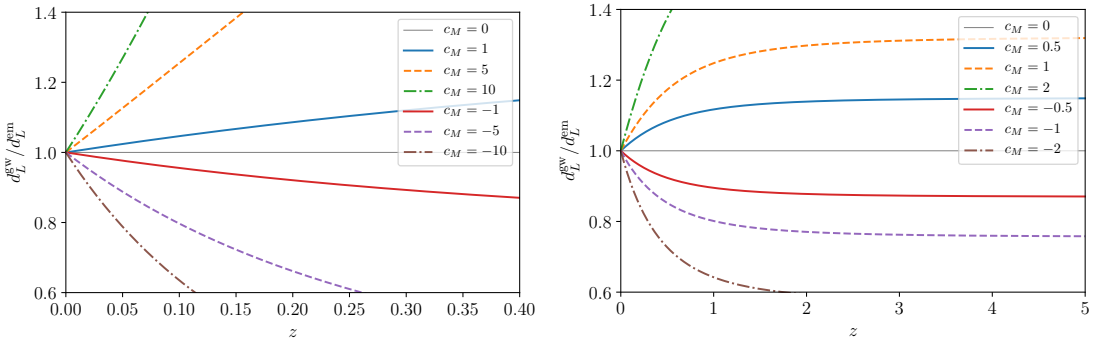

Figure 4.1. Ratio between the GW and the EM luminosity distances due to an additional friction term in the propagation ν scaling with the DE, $\nu(z) = c_{_M}\Omega_{_{DE}}(z)/\Omega_{_{DE}}(0)$. We compare low- (left) and high-redshift (right) signals expected for ground- and space-base GW detectors respectively.

can only travel in the extra dimensions above a certain screening scale R_c, the previous relation generalizes to [391]

$$\frac{d_L^{\rm gw}(z)}{d_L^{\rm em}(z)} = \left[1 + \left(\frac{d_L^{\rm em}}{R_c}\right)^n\right]^{(D-4)/(2n)} , \tag{4.0.4}$$

where n measures the transition steepness and the GR limit is recovered when $D = 4$.

In this chapter, we will focus on changes in the GW luminosity distance due to a modified propagation equation (2.3.2). In section 4.1, we will derive the general formula for $d_L^{\rm gw}$ in terms of ν, the additional friction term from modifying gravity, and c_g, the anomalous speed. For the moment, let us limit to theories in which $c_g = c$. The GW luminosity distance is then related to the EM luminosity distance $d_L^{\rm em}$ by

$$\frac{d_L^{\rm gw}(z)}{d_L^{\rm em}(z)} = \exp\left[\frac{1}{2}\int_0^z \frac{\nu(z')}{1+z'}dz'\right] . \tag{4.0.5}$$

In scalar-tensor gravity, the additional friction is equal to the effective Planck mass run rate $\alpha_{\rm M}$

$$\nu = \alpha_{\rm M} = \frac{d\ln M_*^2}{d\ln a} , \tag{4.0.6}$$

where M_* is the effective Planck mass, i.e. the normalization of the kinetic term of the tensor perturbations. Then, recalling the redshift definition $1 + z = a_0/a$, one arrives at

$$\frac{d_L^{\rm gw}(z)}{d_L^{\rm em}(z)} = \frac{M_*(0)}{M_*(z)} , \tag{4.0.7}$$

where $M_*(0)$ and $M_*(z)$ are the effective Planck masses at the time of observation and emission respectively. At very low redshift, one could assume $\alpha_{\rm M}$ to be constant. Then, our general expression for the ratio of the luminosity distances (4.0.5) simplifies to

$$\frac{d_L^{\rm gw}(z)}{d_L^{\rm em}(z)} = (1+z)^{\alpha_{\rm M}/2} . \tag{4.0.8}$$

A better suited approximation is to assume that the additional friction scales with the DE [392], i.e.

$$\nu(z) = c_{_M}\,\frac{\Omega_{_{DE}}(z)}{\Omega_{_{DE}}(0)}\,, \tag{4.0.9}$$

where $c_{_M}$ is constant. This is a reasonable parametrization for modify gravity theories trying to explain the late time cosmic acceleration. Assuming that the background cosmology is ΛCDM, one obtains

$$\frac{d_L^{\rm gw}}{d_L^{\rm em}} = \exp\left[\frac{1}{2}\frac{c_{_M}}{\Omega_{_{DE}}(0)}\log\left[\frac{1+z}{(\Omega_{_M}(0)(1+z)^3+\Omega_{_{DE}}(0))^{1/3}}\right]\right]. \tag{4.0.10}$$

For illustration, we plot in Fig. 4.1 how the ratio $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ would vary for different values of $c_{_M}$. Positive values of $c_{_M}$ make $d_L^{\rm gw}$ to be larger than $d_L^{\rm em}$ and vice versa. We compare the effects at low redshift (left panel) with high redshift (right panel) expected for ground-base and space-base interferometers respectively. Clearly, low redshift signals are less effective at constraining a modification in the propagation of the GWs. In the case of scalar-tensor gravity, present cosmological survey place order 1 constraints on $c_{_M}$, while future ones could reach $\mathcal{O}(0.1)$ [186].

Because the amount of damping of the signal depends on the distance travelled, this effect would be different for GWs emitted at different redshift. Then, it is necessary multiple detections to constrain this modification of the propagation. In section 4.2 we will discuss how to probe $d_L^{\rm gw}$ using GWs alone and multi-messenger GW events. These methods will also serve to test another type of modification of the GW luminosity distance caused by the mixing of the original GW with additional tensor modes. We will discuss the later in chapter 5.

4.1 Deriving the modified GW luminosity distance

We start the derivation of the GW luminosity distance from the cosmological GW propagation equation (2.3.2) in the absence of sources, $\Pi_{ij} = 0$. This equation can be captured in a standard wave equation,

$$\tilde{\Box} h_{ij} = \left(-\frac{1}{c_g^2}\frac{\partial^2}{\partial\eta^2} + \tilde{\nabla}^2\right) h_{ij} = 0\,, \tag{4.1.1}$$

introducing an effective line element for the GW propagation,

$$d\tilde{s}^2 = a^2(\eta)\,\frac{M_*^2(\eta)}{\bar{M}^2}\,c_g(\eta)\,\left(-c_g^2(\eta)\,d\eta^2 + d\vec{x}^2\right), \tag{4.1.2}$$

where $\bar{M}$ is a normalization scale. This effective line element accounts for the expansion of the universe via $a(\eta)$, the possible modification of the GW light-cone through $c_g(\eta)$ and the additional damping that we parametrize with $\nu = d\ln M_*^2/d\ln a$.[1] In the absence

[1] As mentioned before, in scalar-tensor gravity, M_* can be associated with the effective Planck mass. Here, we do not assign it any specific interpretation and just defined it via ν. This means that our findings will apply to generic modifications of the propagation and not only scalar-tensor theories.

of an anomalous speed and extra friction, this expression reduce to the usual FRW background (2.2.1). As one could note, we have not included an effective graviton mass in (4.1.2). This is because, for astrophysical sources, the frequency of the wave will be large compared to any other parameter and the m_g^2 term in the phase $\omega(k)^2 = c_g^2 k^2 + m_g^2$ will be subdominant.[2]

The simplicity of (4.1.2) allows for an easy understanding of its physical effects. Using purely kinematic arguments, analogous to those discussed in [250] for GR, we will obtain the GW luminosity distance. The first thing to notice is that the standard motion of GWs in null geodesics, $d\tilde{s}^2 = 0$, is only modified by the time dependent tensor speed c_g, since M_* and a are just conformal factors. On the contrary, physical distances $r_{\rm phy}$ will scale with the spatial part of the metric, thus being sensitive also to M_* and a. Traveling thorough null geodesics, the comoving distance between the source and the observer is

$$r_{\rm com}(t) \,=\, |\Delta x| \,=\, \int_{t_s}^{t} \frac{c_T(t')\,dt'}{a(t')}\,, \tag{4.1.3}$$

with t an arbitrary coordinate time between the time of emission t_s and the time of observation t_0, which we assume being today. Differently form GR, the comoving distance depends on the speed of the GW. With this result, we obtain the physical distance

$$r_{\rm phys}(t) \,=\, a(t)\,\frac{M_*(t)}{\bar{M}}\,c_T^{1/2}(t)\,r_{\rm com}(t)\,. \tag{4.1.4}$$

If we demand the GW signal to be as in GR at the time of emission, we have to choose our normalization scale $\bar{M}$ so that in the limit $t \to t_s$ the ratio between physical and comoving distance acquires the standard expression

$$\lim_{t\to t_s} \frac{r_{\rm phys}(t)}{r_{\rm com}(t)} \,=\, a(t_s)\,. \tag{4.1.5}$$

This condition leads to the definition

$$\bar{M} \,\equiv\, M_{\rm eff}(t_s)\,c_T^{1/2}(t_s) \tag{4.1.6}$$

that we will use in this derivation.

Next, we have to derive the frequency-redshift relation. Suppose now that the observer measures GW signals corresponding to wavecrests emitted at different times from the source, which have travelled through the same comoving distance. Using expression (4.1.3) for the comoving distance at the time of observation $r_{\rm com}(t_0)$, at linear order in Δt_s we find the relation

$$\Delta t_0 \,=\, \frac{c_T(t_s)}{c_T(t_0)}\,\frac{a(t_0)}{a(t_s)}\,\Delta t_s \tag{4.1.7}$$

between the time difference of two GW wavecrests as measured at emission and observation times. This formula (4.1.7), which we more conveniently express in terms of

[2]Technically, the mass term enters at higher order in the short-wave approximation. We will present and compare different approximations schemes in section 5.1.1.

redshift, states that source's and observer's clocks tick with different rates. We assume that the observer makes its measurement today at redshift equal to zero, while the the emission occurs at redshift z. Then, we can write

$$dt_0 \;=\; \frac{c_T(z)}{c_T(0)}\;(1+z)\;dt_s\,, \tag{4.1.8}$$

implying that frequencies measured in the source $f^{(s)}$ and observer $f^{(obs)}$ frames are related by

$$f^{(obs)} = \frac{c_g(0)}{c_g(z)}\frac{f^{(s)}}{1+z}\,. \tag{4.1.9}$$

If we call $\mathcal{F}(t)$ the GW energy flux measured by an observer at time t, corresponding to the amount of GW energy per unit time per unit area

$$\mathcal{F}(t) \;\equiv\; \frac{\mathcal{L}(t)}{\text{Area}(t)} = \frac{\mathcal{L}(t)}{4\pi r^2_{\text{phys}}(t)}\,, \tag{4.1.10}$$

and $\mathcal{L}(t)$ the luminosity of the source, defined from the power it radiates at t_s

$$\mathcal{L}(t_s) \;=\; \frac{dE_s}{dt_s}\,, \tag{4.1.11}$$

we can define the luminosity distance d_L^{gw} as

$$d_L^{\text{gw}} = \sqrt{\frac{\mathcal{L}(t_s)}{4\pi\mathcal{F}(t_0)}}\,. \tag{4.1.12}$$

Since we measure the energy flux at the observer position $\mathcal{F}(t_0)$, we need to convert $\mathcal{L}(t_s)$ into the observer frame. The energy scales as the frequency following (4.1.9), while dt_s and dt_0 are related by (4.1.8). Hence, in terms of the redshift,

$$\mathcal{L}(z) = \frac{c_T^2(z)}{c_T^2(0)}\;(1+z)^2\;\mathcal{L}(0)\,. \tag{4.1.13}$$

Using the definition of the physical distance (4.1.4), we obtain

$$d_L^{\text{gw}} = \sqrt{\frac{\mathcal{L}(z)}{\mathcal{L}(0)}}\,r_{\text{phys}}(0) = a(0)\sqrt{\frac{c_T(z)}{c_T(0)}}\,\frac{M_*(0)}{M_*(z)}\,(1+z)\;r_{\text{com}}(0)\,. \tag{4.1.14}$$

If we rewrite the comoving distance in terms of the Hubble expansion,

$$a(0)\,r_{\text{com}}(0) \;=\; \int_0^z \frac{c_g(z')\,dz'}{H(z')}\,, \tag{4.1.15}$$

and M_* in terms of the friction ν,

$$\frac{M_*(0)}{M_*(z)} = \exp\left[\frac{1}{2}\int_0^z \frac{\nu(z')}{1+z'}dz'\right]\,, \tag{4.1.16}$$

we arrive at the final expression

$$d_L^{\rm gw} = \sqrt{\frac{c_T(z)}{c_T(0)}}\,\exp\left[\frac{1}{2}\int_0^z \frac{\nu(z')}{1+z'}dz'\right]\,(1+z)\int_0^z \frac{c_g(z')\,dz'}{H(z')}\,. \tag{4.1.17}$$

This formula for the GW luminosity distance generalizes the expression (4.0.5), which corresponds to the limit $c_g \to c$. Having an anomalous propagation speed, $c_g \neq c$, introduces an extra square root factor, which is nothing but the usual WKB scaling. The propagation speed also affects the integral of the Hubble distance that now cannot be directly linked with the EM luminosity distance.

The amplitude of the GW will then be inversely proportional to $d_L^{\rm gw}$,

$$|h_{+,\times}| = \frac{\tilde{\mathcal{M}}_z^{5/3}(f_{\rm gw}^{\rm (obs)})^{2/3}}{d_L^{\rm gw}}F_{+,\times}\,, \tag{4.1.18}$$

where, in order to express the amplitude in terms of the observed frequency, we have redefined the redshifted chirp mass to

$$\tilde{\mathcal{M}}_z \equiv \frac{c_g(z)}{c_g(0)}(1+z)\mathcal{M}_c\,. \tag{4.1.19}$$

Several comments are in order. First of all, we note that the luminosity distance (4.1.17) and the redshifted chirp mass (4.1.19) depend on the anomalous speed $\alpha_{_T} \equiv c_g^2 - 1$ and friction $\nu = d\log(M_*^2)/d\log(a)$. This implies that the parameters of the binary, e.g. the masses of the compact objects, could be different to the ones obtained assuming ΛCDM. However, due to the stringent constraints in the speed of GWs after GW170817, see (3.2.4), we have effectively $c_g/c = 1$. Therefore, this degeneracy between the gravitational sector and the compact binary characterization is broken and the usual results are recovered.

4.2 Probing the damping of GWs

As discussed before, a modification of the GW luminosity distance can affect different GW observables. In the following we are going to define two broad classes of tests to probe this modification. First we will study those tests with GW signals alone. Then we will consider multi-messenger tests.

GWs tests: with GWs alone we could use the merger rate of compact binaries as a function of redshift $R(z)$. If there is an additional friction term $\nu(z)$, the amplitude of the GWs at the detector will be lower and, as a consequence, less events would be detected. Comparing the observed rate with the theoretical prediction from a given astrophysical model could constrain this modification in the propagation. Forecasts of the possible constraints using GW number counts have been done in Ref. [393]. The positive side of this type of test is that it does not require a redshift identification. Therefore, one could use BBH directly. This is convenient in two ways: *i)* BBH events are more numerous

and *ii)* BBH can be detected at higher redshift. As we have seen in Fig. 4.1, the effect of an absorptive medium is more significant at high-z. This is specially promising for third-generation GW detectors which could hear BBH up to $z \sim 10 - 15$. On the negative side, there is an intrinsic uncertainty in the theoretical modeling of the merger rate that might limit the precision. However, if the merger rate exhibits a particular pattern with redshift, it could be easy to disentangle the degeneracy with astrophysics. For instance, in chapter 5, we will study the phenomenon of GW oscillations. This can lead to an oscillatory behavior of the merger rate that is potentially distinguishable. We will analyze this scenario in section 5.2.2.

Multimessenger tests: with multi-messenger events, we can constrain this effect further. Either by a direct EM counterpart or a statistical analysis, if we can determine the EM luminosity distance of the source, we can then test the ratio $d_L^{\rm gw}/d_L^{\rm em}$. Any deviation of this ratio from being 1 would be a smoking gun for physics beyond GR in the standard model of cosmology ΛCDM.

With the detection of the multi-messenger event GW170817 it was possible to test the gravitational Hubble diagram for the first time. The observation was consistent with GR although being just one event the constraining power is still moderate. For theories with extra dimensions following (4.0.3), it was found that the number of space-time dimensions in which the gravitons propagate is limited to [394]

$$D = 4.02^{+0.07}_{-0.10} \quad \text{or} \quad D = 3.98^{+0.07}_{-0.09} \tag{4.2.1}$$

for SN or CMB prior in H_0 (see Fig. 2.3 and Sec. 2.2.1).[3] Nevertheless, the additional friction ν can only be loosely constrained [379]

$$-75.3 \leqslant \nu(0) \leqslant 78.4\,. \tag{4.2.2}$$

An important remark when evaluating the GW luminosity distance in modified gravity is that it will not only be altered with respect to GR due to the modified propagation of GWs but also because the cosmological expansion history is different. In other words, in alternative theories of gravity both the EM luminosity distance $d_L^{\rm em}$ and its relation with the GW luminosity distance can be modified, due to a different $H(z)$ and to an additional friction ν respectively. In fact, the contribution of the the modified propagation can dominate over the modified cosmic expansion history. Introducing a phenomenological parametrization of the GW luminosity distance [395]

$$\frac{d_L^{\rm gw}(z)}{d_L^{\rm em}(z)} = \Xi_0 + \frac{1-\Xi_0}{(1+z)^n} \tag{4.2.3}$$

together with the usual (w_0, w_a) parametrization of $H(z)$, it was shown that the largest contributions are Ξ_0 and w_0. In [8], we studied in detail how this parametrization adjusts to several modified gravity theories. From Fig. 4.1 one can already anticipate that this

[3] This model was reanalyzed in [391] without assuming any prior in H_0 but the GW170817 measurement and including the screening (4.0.4), which differs from the one in [394].

parametrization captures accurately the plateau at high-redshift. It has however more limitations at low redshift but this is only due to its simplicity. We also investigated the prospects of measuring Ξ_0 with LISA using as standard sirens SMBHs. We found that Ξ_0 could be constrained to the $1-5\%$ level depending on the assumed number of detections and astrophysical details. This shows the potential of LISA to probe several modify gravity models.

CHAPTER

5

GW oscillations

GR predicts that there are only two propagating degrees of freedom: the $h_{+,\times}$ polarizations. Extensions of GR can however have more propagating modes which in general will mix along the propagation. We will study the propagation of these perturbations on top of homogeneous and isotropic, cosmological backgrounds. Depending on the concrete model, this Cosmological Principle will be realised distinctively built upon different symmetries owed to different field configurations. For now we will not yet particularize to concrete theories but consider an effective parametrization of the underlying physics. Compatible with the symmetries of Friedmann-Lemaître-Robertson-Walker (FLRW) space-times, we will assume the background to be of the form $ds^2 = a^2(-d\eta^2+\delta_{ij}dx^i dx^j)$ in conformal time $d\eta = dt/a^2$. On top of these homogeneous and isotropic backgrounds we will consider small perturbations. They can generally be decomposed in terms of scalar, vector, and tensor perturbations as directly arising from irreducible representations of the $SO(3)$ background field configuration.

One immediate property of the background symmetries is that the different sectors of perturbations decouple from each other. We will be specially interested in the tensor perturbations $g_{ij} = \bar{g}_{ij}^{\rm FLRW} + h_{ij}$, where they are functions of time and space and accord to the transformations under spatial rotations. Importantly, the tensor perturbations can only couple to other tensor sectors but are completely decoupled from the scalar and vector perturbations at the linear order. Therefore, to study the possible mixing of $h_{+,\times}$ with non-GR modes, we will be interested in alternative theories containing additional tensor modes. These extra tensors could either originate from a second spin-2 field $f_{\mu\nu}$ or a set of vector fields A^a_μ. We will denote them generically by t_{ij}. This enables the possibility of oscillations between h_{ij} and t_{ij} if they are coupled.

One should note that GW oscillations could appear in other contexts. For example, if the background is not FRW, GR polarizations could mix with scalar and vector modes. Although studying the propagation of GWs over non-cosmological background goes beyond the scope of this chapter, let us emphasis that the framework that we are going to develop could be easily extended.

5.1 General formalism

In this section we are going to introduce a general framework to study GW oscillations produced during the propagation of different tensor perturbations over cosmological

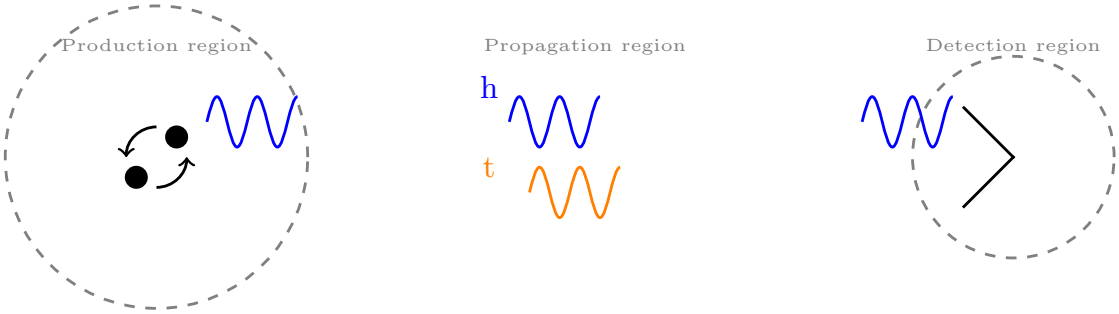

Figure 5.1. Schematic representation of the different regions between the source and the detector: production, propagation, detection. We assume that *(i)* GWs are generated as in GR, *(ii)* the second tensor t is only excited in the propagation region and *(iii)* only h couples to matter and, thus, the detector.

backgrounds. In order to connect with observations, we will consider the following assumptions throughout the chapter:

(i) only one of the perturbations, $h_{\mu\nu}$, interact with matter;

(ii) the production of GWs follows that of GR, as supported by the decay of the orbit of binary pulsars [305], at least in the region in which the post-Newtonian expansion holds;[1]

(iii) the production and detection regions are small compared to the propagation zone (see Fig. 5.1) so that we can consider the GW propagating over the cosmological background from emission to detection;

(iv) there are not significant deviations of the cosmological background in the propagation zone.

Therefore, the main observable will be the transfer function of the amplitude $T(\eta, k)$ and the phase of the wave $\theta(\eta, k)$,

$$h_{+,\times}(\eta, k) = h^{\rm GR}_{+,\times}(\eta, k) \cdot |T_{+,\times}(\eta, k)| e^{i \int \theta_{+,\times}(\eta,k) d\eta} \,, \tag{5.1.1}$$

where $h^{\rm GR}$ denotes the GR signal. One should notice that, although in this analysis we are neglecting any modification in the emission or in the cosmological background, those effects could be incorporated by complicating the schema of zones presented. For instance, if the emission is modified, one only needs to take the appropriate function as the initial condition of the propagation region. On the other hand, if there is a region in which the background is not FLRW, one would need to add an additional transfer function in this new zone.

[1] This assumption will break near the merger, where the GW scattering is large (disregarding other possible effects such as absorption, dispersion and diffraction). The modifications in the production of GWs go out of scope of this work and generally needs highly involved numerical analysis.

At this point it is worthwhile to mention that in the following we are going to solve the evolution of monochromatic waves. In reality, compact binary mergers produce wave packets of a given duration. In the case in which there is a modified dispersion relation and the frequency of the wave changes rapidly (near the merger for instance), there could be interference within the wave-packet. This could lead to new observational effects. In the inspiral part however, we expect these corrections to be small.

5.1.1 Solving the evolution: WKB v.s. large-k expansion

Without specifying the underlying covariant theory, we can write down the most general coupled equations of motion on top of a cosmological background and work in terms of these parametrized quantities. The evolution of the linear tensor perturbations $h_{\mu\nu}$ and $t_{\mu\nu}$ forms a system of coupled, second order differential equations. For the subsequent discussion, it will be convenient to present it in matrix notation as

$$\left[\frac{d^2}{d\eta^2} + \hat{\nu}\frac{d}{d\eta} + \hat{C}k^2 + \hat{\Pi}k + \hat{M}\right]\begin{pmatrix} h \\ t \end{pmatrix} = 0\,, \tag{5.1.2}$$

where we have defined the friction matrix $\hat{\nu}$, the velocity matrix $\hat{C}$, the chirality matrix $\hat{\Pi}$, the mass matrix $\hat{M}$, and we are measuring in conformal time η.[2] One should note that in the above expression we have dropped the indices because the equations are the same for the two transverse, traceless polarizations. Whenever there are not parity violating terms $\hat{\Pi}$, we will be implicitly working in the usual $+$ and $\times$ basis (although the equation will not change for L and R polarizations). On the contrary, when $\hat{\Pi} \neq 0$, we will refer to the circular polarizations left L and right R because in this case the polarization basis matters.

In general, any of these matrices can be non-diagonal and thus trigger a mixing of the two modes. Moreover, the entries of these matrices are also generically time-dependent. As a consequence, there will be no exact, analytic solutions. For that reason, in order to understand the physics of the problem, we present different schemes to obtain approximate solutions: one based on a Wentzel-Kramers-Brillouin (WKB) expansion and another on a large wavenumber k expansion.

Concretizing down to a model would mean fixing the coefficients of these matrices in a specific way in terms of the background evolution. In some cases, some of the entries will be even associated with each other. The aim will be to break the degeneracies between the parameters using the full-fledged observational information. In most cases, this will require the combination of various observational channels.

WKB expansion

We are interested in theories modifying gravity at cosmological scales. Therefore, the typical time variation of the parameters will be of order of the inverse Hubble constant

[2] Note that if the fields h and t are normalized canonically and if the equations (5.1.2) come from a Lagrangian, there are relations between different matrix elements of $\hat{\nu}$, $\hat{C}$, $\hat{\Pi}$ and $\hat{M}$.

H_0. On the other hand, the frequency of a GW from a compact binary merger scales as

$$f_{\rm gw} \sim 1{\rm kHz}\left(\frac{10M_\odot}{M}\right) \sim 10^{21} H_0 \left(\frac{10M_\odot}{M}\right) . \tag{5.1.3}$$

Therefore, for any signal of this kind there will be a great difference between the time scales of the problem, having $f_{\rm gw} \gg H_0$. This motivates solving equation (5.1.2) using an adiabatic or WKB approximation. For that, we introduce a dimensionless, small parameter ϵ suppressing the time derivatives

$$\left[\epsilon^2 \frac{d^2}{d\eta^2} + \epsilon\, \hat{\nu} \frac{d}{d\eta} + \hat{C}k^2 + \hat{\Pi}k + \hat{M}\right] \vec{\Phi} = 0 \tag{5.1.4}$$

and enhancing the phase of the wave

$$\vec{\Phi} = \hat{E} e^{\frac{i}{\epsilon}\int \hat{\theta} d\eta} \left(\vec{\Phi}_0 + \epsilon \vec{\Phi}_1 + \cdots\right) , \tag{5.1.5}$$

where $\vec{\Phi}$ stands for

$$\vec{\Phi} = \begin{pmatrix} h \\ t \end{pmatrix} . \tag{5.1.6}$$

Since we are in a multidimensional problem, we are expanding the solution around the basis determined by the matrix $\hat{E}$ solving the constant-parameter case. The amplitude is expanded in different orders of ϵ and $\hat{\theta}$ is the diagonal phase matrix. Defining $\hat{G} \equiv e^{\frac{i}{\epsilon}\int \hat{\theta} d\eta}$, then we have the following equations at increasing order in ϵ

$$\epsilon^0 : \qquad \left(\left(\hat{C}k^2 + \hat{\Pi}k + \hat{M}\right)\hat{E} - \hat{E}\hat{\theta}^2 + i\hat{\nu}\hat{E}\hat{\theta}\right)\hat{G}\vec{\Phi}_0 = 0 , \tag{5.1.7}$$

$$\epsilon^1 : \qquad \left(2\hat{E}\hat{\theta} - i\hat{\nu}\hat{E}\right)\hat{G}\vec{\Phi}_0' + \left(\hat{E}\hat{\theta}' + 2\hat{E}'\hat{\theta} - i\hat{\nu}\hat{E}'\right)\hat{G}\vec{\Phi}_0 = 0 , \tag{5.1.8}$$

$$\epsilon^2 : \qquad \hat{E}\hat{G}\vec{\Phi}_0'' + 2\hat{E}'\hat{G}\vec{\Phi}_0' + \hat{E}''\hat{G}\vec{\Phi}_0 = \tag{5.1.9}$$
$$- i\left(2\hat{E}\hat{\theta} - i\hat{\nu}\hat{E}\right)\hat{G}\vec{\Phi}_1' - i\left(\hat{E}\hat{\theta}' + 2\hat{E}'\hat{\theta} - i\hat{\nu}\hat{E}'\right)\hat{G}\vec{\Phi}_1 ,$$

$$\cdots \qquad \cdots$$

$$\epsilon^{n+1} : \qquad \hat{E}\hat{G}\vec{\Phi}_{n-1}'' + 2\hat{E}'\hat{G}\vec{\Phi}_{n-1}' + \hat{E}''\hat{G}\vec{\Phi}_{n-1} = \tag{5.1.10}$$
$$- i\left(2\hat{E}\hat{\theta} - i\hat{\nu}\hat{E}\right)\hat{G}\vec{\Phi}_n' - i\left(\hat{E}\hat{\theta}' + 2\hat{E}'\hat{\theta} - i\hat{\nu}\hat{E}'\right)\hat{G}\vec{\Phi}_n .$$

To solve the leading order equation, which gives the exact solution when the coefficients are constant, we have to find the roots of the quartic equation

$$\det\left[\hat{C}k^2 + \hat{\Pi}k + \hat{M} - \hat{I}\theta^2 + i\hat{\nu}\theta\right] = 0 . \tag{5.1.11}$$

The matrix $\hat{E}$ is then

$$\hat{E} = \begin{pmatrix} 1 & -\frac{\hat{W}_{12} + i\hat{\nu}_{12}\theta_i}{\hat{W}_{11} - \theta_i^2 + i\hat{\nu}_{11}\theta_i} \\ -\frac{\hat{W}_{21} + i\hat{\nu}_{21}\theta_j}{\hat{W}_{22} - \theta_j^2 + i\hat{\nu}_{22}\theta_j} & 1 \end{pmatrix} , \tag{5.1.12}$$

where for shortness we have defined $\hat{W} \equiv \hat{C}k^2 + \hat{\Pi}k + \hat{M}$. Note, that since we have dropped the indices for the two polarizations, the matrices are all 2×2 matrices. Here, θ_i and θ_j correspond to two different solutions of Eq. (5.1.11).

At next to leading order, $\mathcal{O}(\epsilon^1)$, we solve $\vec{\Phi}_0$ from the first order differential equation (5.1.8). In general, this is a system of first order ordinary differential equations with time dependent coefficients without analytic solutions.[3] However, within the WKB expansion the matrix exponential is a good approximate solution.[4] Accordingly, we can solve $\vec{\Phi}_0$ as

$$\vec{\Phi}_0 = \hat{\theta}^{-1/2} e^{-\int \hat{A}_{\rm wkb}\, d\eta}\, \vec{C}_0\,, \tag{5.1.13}$$

where the matrix in the exponent corresponds to

$$\hat{A}_{\rm wkb} = \hat{G}^{-1}\hat{\theta}^{1/2}\left(2\hat{E}\hat{\theta} - i\hat{\nu}\hat{E}\right)^{-1}\left(2\hat{E}'\hat{\theta} - i\hat{\nu}\hat{E}' + \frac{i}{2}\hat{\nu}\hat{E}\hat{\theta}'\hat{\theta}^{-1}\right)\hat{G}\hat{\theta}^{-1/2}\,, \tag{5.1.14}$$

and $\vec{C}_0$ is a vector of constant coefficients to be fixed with the initial conditions. Here one should recall that $\hat{\theta}$ is a diagonal matrix and thus the term $\hat{\theta}^{-1/2}$ in front is just the usual WKB scaling $1/\sqrt{\theta_i}$ of the one-dimensional problem. If there is time dependence, there can be corrections to this scaling, which corresponds to the matrix exponential.

At next to next to leading order, the first correction to the amplitude $\vec{\Phi}_1$ can be computed from (5.1.9), which is analogous to (5.1.8) but with a non-homogeneous term. In fact, the solution of the n-th correction will have the same structure given by the iterative solution

$$\vec{\Phi}_n = \hat{\theta}^{-1/2} e^{-\int \hat{A}_{\rm wkb} d\eta}\left(\vec{C}_n + i\int e^{\int \hat{A}_{\rm wkb} d\eta}\, \hat{B}^{-1}_{\rm wkb}\, \vec{F}^{\rm wkb}_{n-1}\, d\eta\right)\,, \tag{5.1.15}$$

where

$$\hat{B}_{\rm wkb} = \left(2\hat{E}\hat{\theta} - i\hat{\nu}\hat{E}\right)\hat{G}\hat{\theta}^{-1/2}\,, \tag{5.1.16}$$

$$\vec{F}^{\rm wkb}_{n-1} = \left(\hat{E}\hat{G}\vec{\Phi}''_{n-1} + 2\hat{E}'\hat{G}\vec{\Phi}'_{n-1} + \hat{E}''\hat{G}\vec{\Phi}_{n-1}\right)\,, \tag{5.1.17}$$

and $\vec{C}_n$ is a constant vector. In this way we have solved the problem up to order ϵ^{n+1}.

The above general solution can be simplified in some cases. For instance, when $\hat{\nu}$ commutes with $\hat{E}$ and $\hat{\theta}$, the friction matrix $\hat{\nu}$ may be absorbed by defining

$$\vec{\Phi}\Big|_{[\hat{E},\hat{\nu}]=0} = e^{-\frac{1}{2}\int \hat{\nu} d\eta}\hat{E}e^{\frac{i}{\epsilon}\int \hat{\theta} d\eta}\left(\vec{\Phi}_0 + \epsilon\vec{\Phi}_1 + \cdots\right)\,, \tag{5.1.18}$$

where again we are using the matrix exponential as an approximate solution whenever $\hat{\nu}$ is non-diagonal. Then (5.1.11) becomes a quadratic equation for θ^2,

$$\theta^2_{1,2} = \frac{1}{2}\left(\mathrm{tr}\left\{\hat{W}\right\} \pm \sqrt{4\hat{W}_{12}\hat{W}_{21} + \left(\hat{W}_{11} - \hat{W}_{22}\right)^2}\right)\,, \tag{5.1.19}$$

[3] Although a *formal* analytical solution could be written in terms of time-ordered exponential.

[4] To be an exact solution, the matrix in the exponent should commute with itself at any two instants of time. Because the parameters vary slowly in the WKB compared to the GW frequency this is a good approximation.

and $\hat{E}$ is the associated matrix of eigenvectors of $\hat{W}$.

Altogether, we can decompose the general solution in each of the components, obtaining

$$h(\eta) = \left[c_1\Phi_h(\eta) + c_2\hat{E}_{12}(\eta)\Phi_t(\eta)e^{i\int\delta\theta(\eta)d\eta}\right]e^{i\int\theta_1(\eta)d\eta}\,, \tag{5.1.20}$$

$$t(\eta) = \left[c_2\Phi_t(\eta) + c_1\hat{E}_{21}(\eta)\Phi_h(\eta)e^{-i\int\delta\theta(\eta)d\eta}\right]e^{i\int\theta_2(\eta)d\eta}\,. \tag{5.1.21}$$

Here, we have denoted the difference in the phases as $\delta\theta = \theta_2 - \theta_1$, and $\Phi_{h,t}$ incorporate all the corrections from $\sum_n \vec{\Phi}_n$ to the amplitude of h and t respectively. One should note that the above expressions (5.1.20-5.1.54) corresponds only to the contribution of two distinct (in absolute value) phases $\theta_{1,2}$. Whenever there are four independent roots of (5.1.11), one should add to (5.1.20-5.1.54) the equivalent terms depending on $\theta_{3,4}$. Finally, we can fix the constants $c_{1,2}$ using the initial conditions at the time of emission η_e. Imposing that initially only one of the tensor perturbations is excited with an amplitude h_0 dictated by GR, i.e. $h(\eta_e) = h_0$ and $t(\eta_e) = 0$, we find

$$c_1 = \frac{h_0}{\Phi_h(\eta_e)(1 - \hat{E}_{12}(\eta_e)\hat{E}_{21}(\eta_e))}\,, \tag{5.1.22}$$

$$c_2 = -\frac{h_0\hat{E}_{21}(\eta_e)}{\Phi_h(\eta_e)(1 - \hat{E}_{12}(\eta_e)\hat{E}_{21}(\eta_e))}\,. \tag{5.1.23}$$

Large-k expansion

In addition to the hierarchy between the time variation of the parameters of the theory and the frequency of the GWs, it could be the case that the parameters themselves are small compared to the wavenumber k. Accordingly, one could make a large-k or shortwave expansion (also known as eikonal approximation [249]), which is a more restrictive approximation compared to the WKB. Using the same ansatz for $\vec{\Phi}$, the system of equations is however different, i.e.

$$\left[\frac{d^2}{d\eta^2} + \hat{\nu}\frac{d}{d\eta} + \epsilon^{-2}\hat{C}k^2 + \epsilon^{-1}\hat{\Pi}k + \hat{M}\right]\vec{\Phi} = 0. \tag{5.1.24}$$

Splitting in the different orders, we find

$$\epsilon^{-2}: \qquad \left(\hat{C}\hat{E}k^2 - \hat{E}\hat{\theta}^2\right)\hat{G}\vec{\Phi}_0 = 0\,, \tag{5.1.25}$$

$$\epsilon^{-1}: \qquad 2\hat{E}\hat{\theta}\hat{G}\vec{\Phi}_0' + \left(\hat{E}\hat{\theta}' + 2\hat{E}'\hat{\theta} + \hat{\nu}\hat{E}\hat{\theta} - i\hat{\Pi}\hat{E}k\right)\hat{G}\vec{\Phi}_0 = 0\,, \tag{5.1.26}$$

$$\epsilon^{0}: \qquad \hat{E}\hat{G}\vec{\Phi}_0'' + \left(2\hat{E}' + \hat{\nu}\hat{E}\right)\hat{G}\vec{\Phi}_0' + \left(\hat{E}'' + \hat{\nu}\hat{E}' + \hat{M}\hat{E}\right)\hat{G}\vec{\Phi}_0 = \tag{5.1.27}$$
$$- 2i\hat{E}\hat{\theta}\hat{G}\vec{\Phi}_1' - i\left(\hat{E}\hat{\theta}' + 2\hat{E}'\hat{\theta} + \hat{\nu}\hat{E}\hat{\theta} - i\hat{\Pi}\hat{E}k\right)\hat{G}\vec{\Phi}_1\,,$$

$$\cdots \qquad \cdots$$

$$\epsilon^{n-1}: \qquad \hat{E}\hat{G}\vec{\Phi}_{n-1}'' + \left(2\hat{E}' + \hat{\nu}\hat{E}\right)\hat{G}\vec{\Phi}_{n-1}' + \left(\hat{E}'' + \hat{\nu}\hat{E}' + \hat{M}\hat{E}\right)\hat{G}\vec{\Phi}_{n-1} = \tag{5.1.28}$$
$$- 2i\hat{E}\hat{\theta}\hat{G}\vec{\Phi}_n' - i\left(\hat{E}\hat{\theta}' + 2\hat{E}'\hat{\theta} + \hat{\nu}\hat{E}\hat{\theta} - i\hat{\Pi}\hat{E}k\right)\hat{G}\vec{\Phi}_n\,.$$

To solve the leading order equation, we take θ^2 as the eigenvalues of $\hat{C}$ (cf. (5.1.19)) and $\hat{E}$ the matrix of eigenvectors

$$\hat{E} = \begin{pmatrix} 1 & -\frac{\hat{C}_{12}}{\hat{C}_{11}-\theta_2^2} \\ -\frac{\hat{C}_{21}}{\hat{C}_{22}-\theta_1^2} & 1 \end{pmatrix} . \tag{5.1.29}$$

In the case in which the velocity matrix $\hat{C}$ is diagonal, which is the most common case, then the matrix of eigenvectors becomes the identity matrix $\hat{E} = \hat{I}$.

At next order, we obtain the amplitude again using an approximate matrix exponential solution

$$\vec{\Phi}_0 = \hat{\theta}^{-1/2} e^{-\int \hat{A}_{\text{large-k}}\, d\eta}\, \vec{C}_0 \,, \tag{5.1.30}$$

but now with a matrix in the exponent

$$\hat{A}_{\text{large-k}} = \frac{1}{2}\hat{G}^{-1}\hat{\theta}^{-1/2}\hat{E}^{-1}\left(2\hat{E}' + \hat{\nu}\hat{E} - i\hat{\Pi}\hat{E}\hat{\theta}^{-1}k\right)\hat{\theta}^{1/2}\hat{G} \tag{5.1.31}$$

different to (5.1.14). The higher-order corrections to the amplitude can be computed as well as before,

$$\vec{\Phi}_n = \hat{\theta}^{-1/2} e^{-\int \hat{A}_{\text{large-k}} d\eta}\left(\vec{C}_n + i\int e^{\int \hat{A}_{\text{large-k}} d\eta}\, \hat{B}^{-1}_{\text{large-k}}\, \vec{F}^{\text{large-k}}_{n-1}\, d\eta\right) , \tag{5.1.32}$$

with

$$\hat{B}_{\text{large-k}} = 2\hat{E}\hat{G}\hat{\theta}^{1/2} \,, \tag{5.1.33}$$

$$\vec{F}^{\text{large-k}}_{n-1} = \hat{E}\hat{G}\vec{\Phi}''_{n-1} + \left(2\hat{E}' + \hat{\nu}\hat{E}\right)\hat{G}\vec{\Phi}'_{n-1} + \left(\hat{E}'' + \hat{\nu}\hat{E}' + \hat{M}\hat{E}\right)\hat{G}\vec{\Phi}_{n-1} \,, \tag{5.1.34}$$

In this way we have solved the problem up to order ϵ^{n-1}.

If we focus in the leading order amplitude, we could rewrite the previous formula (5.1.30) as

$$\vec{\Phi}_0 = \hat{\theta}^{-1/2}\, e^{-\frac{1}{2}\,\mathrm{tr}\{\bar{A}\}} \begin{pmatrix} \cos\omega + \frac{\Delta\bar{A}}{2\omega}\sin\omega & -\frac{\bar{A}_{12}}{\omega}\sin\omega \\ -\frac{\bar{A}_{21}}{\omega}\sin\omega & \cos\omega - \frac{\Delta\bar{A}}{2\omega}\sin\omega \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \tag{5.1.35}$$

by denoting the integral of the matrix in the exponent $\bar{A}_{ij} = \int_{\eta_e}^{\eta} \hat{A}_{ij} d\eta$, defining the difference of the diagonal entries $\Delta\bar{A} = \bar{A}_{22} - \bar{A}_{11}$ and introducing a frequency

$$\omega^2 = -\bar{A}_{12}\bar{A}_{21} - \Delta\bar{A}^2/4 \,. \tag{5.1.36}$$

In the case in which the velocity matrix is diagonal, the mixing of the modes is controlled by ω. This is explicit when we compute each tensor perturbation

$$h(\eta) = \frac{e^{-\frac{1}{2}\,\mathrm{tr}\{\bar{A}\}}}{\sqrt{\theta_1}}\left[c_1\left(\cos\omega + \frac{\Delta\bar{A}}{2\omega}\sin\omega\right) - c_2\frac{\bar{A}_{12}}{\omega}\sin\omega\right] e^{i\int\theta_1(\eta)d\eta} \,, \tag{5.1.37}$$

$$t(\eta) = \frac{e^{-\frac{1}{2}\,\mathrm{tr}\{\bar{A}\}}}{\sqrt{\theta_2}}\left[c_2\left(\cos\omega - \frac{\Delta\bar{A}}{2\omega}\sin\omega\right) - c_1\frac{\bar{A}_{21}}{\omega}\sin\omega\right] e^{i\int\theta_2(\eta)d\eta} \,. \tag{5.1.38}$$

when we impose the initial conditions $h(\eta_e) = h_0$ and $t(\eta_e) = 0$, this expression simplifies further to (note that $\bar{A}_{ij}(\eta_e) = 0$)

$$h(\eta) = h_0\, e^{-\frac{1}{2}\,\mathrm{tr}\{\bar{A}\}} \left(\cos\omega + \frac{\Delta\bar{A}}{2\omega}\sin\omega\right) \frac{\sqrt{\theta_1(\eta_e)}}{\sqrt{\theta_1(\eta)}}\, e^{i\int\theta_1(\eta)d\eta}\,, \tag{5.1.39}$$

$$t(\eta) = -h_0\, e^{-\frac{1}{2}\,\mathrm{tr}\{\bar{A}\}} \frac{\bar{A}_{21}}{\omega}\sin\omega \frac{\sqrt{\theta_1(\eta_e)}}{\sqrt{\theta_2(\eta)}}\, e^{i\int\theta_2(\eta)d\eta}\,. \tag{5.1.40}$$

From this is expression we can also see that there will be an overall damping determined by $\mathrm{tr}\Big\{\bar{A}\Big\}$.

5.1.2 Different types of GW mixings

In order to gain insights from the general, approximate, analytical solutions that we have found, let us consider some particular cases. It is important to note that in general there will be degeneracies between different parameters. For this reason, we also consider representative examples separately.

Mixing through the mass matrix

In analogy with neutrinos, if the mass matrix of the tensor perturbations is non-diagonal, the propagation and mass eigenstates are different, implying that they will mix while traveling. In the following we consider h and t propagating at different speeds and interacting through the mass matrix $\hat{M}$,[5]

$$\left[\frac{d^2}{d\eta^2} + \begin{pmatrix} c_h^2 & 0 \\ 0 & c_t^2 \end{pmatrix} k^2 + \begin{pmatrix} m_h^2 & m_{ht}^2 \\ m_{th}^2 & m_t^2 \end{pmatrix}\right] \begin{pmatrix} h \\ t \end{pmatrix} = 0\,. \tag{5.1.41}$$

The associated eigenvalues are

$$\theta_{1,2}^2 = \left(c_h^2 + \frac{1}{2}\Delta c^2\right) k^2 + \frac{1}{2}M^2 \mp \frac{1}{2}\sqrt{\Delta c^4 k^4 + 2\Delta c^2 \Delta m^2 k^2 + M^4(1+\Delta)^2}\,, \tag{5.1.42}$$

where we have defined the difference in the speeds $\Delta c^2 \equiv c_t^2 - c_h^2$, the sum of the square masses $M^2 \equiv m_h^2 + m_t^2$, their difference $\Delta m^2 = m_t^2 - m_h^2$ and the parameter

$$\Delta \equiv \sqrt{1 - 4\det\Big[\hat{M}\Big]/M^4} - 1\,, \tag{5.1.43}$$

which vanishes when $\hat{M}$ is degenerate. The eigenvectors are

$$\hat{E} = \begin{pmatrix} 1 & -\frac{m_{ht}^2}{c_h^2 k^2 + m_h^2 - \theta_2^2} \\ -\frac{m_{th}^2}{c_t^2 k^2 + m_t^2 - \theta_1^2} & 1 \end{pmatrix}\,. \tag{5.1.44}$$

[5] One should note that if the fields are normalized canonically and we assume this equations of motion come from a Lagrangian then $m_{ht} = m_{th}$ and all the expressions in the following discussion simplify.

High-k limit: It is interesting to study first the high-k limit. The phases are

$$\theta_1^2 = c_h^2 k^2 + m_h^2 + \mathcal{O}(k^{-2})\,, \tag{5.1.45}$$
$$\theta_2^2 = (c_h^2 + \Delta c^2)k^2 + m_t^2 + \mathcal{O}(k^{-2})\,, \tag{5.1.46}$$

and the matrix of eigenvectors scales as

$$\hat{E} = \begin{pmatrix} 1 & \frac{m_{ht}^2}{\Delta c^2 k^2 + \Delta m^2 + \mathcal{O}(k^{-2})} \\ -\frac{m_{th}^2}{\Delta c^2 k^2 + \Delta m^2 + \mathcal{O}(k^{-2})} & 1 \end{pmatrix} = \hat{I} + \mathcal{O}\left(\frac{m_{ij}^2}{\Delta c^2 k^2}\right). \tag{5.1.47}$$

Therefore, if h and t propagate at different speeds, $\Delta c \neq 0$, and the wavenumber k is much larger than the matrix elements of $\hat{M}$, the mixing will be suppressed, with $\hat{E}$ approaching the identity matrix.

Small-Δc limit: Since we are interested in studying the regime in which k is large, let us consider the limit in which the difference in the speeds is small, $\Delta c^2 \ll 1$, and the mixing is not suppressed. In this limit, the phases simplify to

$$\theta_1^2 = \left(c_h^2 + \frac{1}{2}\Delta c^2 \left(1 - \frac{\Delta m^2}{M^2(1+\Delta)}\right)\right) k^2 - \frac{1}{2}M^2\Delta + \mathcal{O}(\Delta c^4)\,, \tag{5.1.48}$$
$$\theta_2^2 = \left(c_h^2 + \frac{1}{2}\Delta c^2 \left(1 + \frac{\Delta m^2}{M^2(1+\Delta)}\right)\right) k^2 + M^2(1 + \frac{1}{2}\Delta) + \mathcal{O}(\Delta c^4)\,, \tag{5.1.49}$$

and the eigenvectors to

$$\hat{E} = \begin{pmatrix} 1 & \frac{m_{ht}^2}{m_t^2 + M^2\Delta/2}\left(1 - \frac{\Delta c^2 k^2}{M^2(1+\Delta)} + \mathcal{O}(\Delta c^4)\right) \\ -\frac{m_{th}^2}{m_t^2 + M^2\Delta/2}\left(1 - \frac{\Delta c^2 k^2}{M^2(1+\Delta)} + \mathcal{O}(\Delta c^4)\right) & 1 \end{pmatrix}. \tag{5.1.50}$$

The frequency of oscillation due to the mixing is governed by the difference in the eigenfrequencies

$$\delta\theta \equiv \theta_2 - \theta_1 = \frac{M^2(1+\Delta)}{2c_h k} + \mathcal{O}(k^{-3}, \Delta c^2)\,, \tag{5.1.51}$$

which suggest the introduction of the effective mass

$$m_g \equiv M\sqrt{1+\Delta}\,. \tag{5.1.52}$$

In the same manner, whenever the components of the mass matrix scale similarly with time, e.g. $m_{ij}^2 \propto a(\eta)^2$, the matrix of eigenvectors becomes approximately constant, i.e. $\hat{E} = \text{const.} + \mathcal{O}(\Delta \mathrm{c}^2 \mathrm{k}^2/\mathrm{M}^2)$. The WKB solutions are then

$$h(\eta) = \left[\frac{c_1}{\sqrt{\theta_1(\eta)}} + \frac{c_2}{\sqrt{\theta_2(\eta)}}\hat{E}_{12}(\eta) e^{i\int \delta\theta(\eta)d\eta}\right] e^{i\int \theta_1(\eta)d\eta}\,, \tag{5.1.53}$$
$$t(\eta) = \left[\frac{c_2}{\sqrt{\theta_2(\eta)}} + \frac{c_1}{\sqrt{\theta_1(\eta)}}\hat{E}_{21}(\eta) e^{-i\int \delta\theta(\eta)d\eta}\right] e^{i\int \theta_2(\eta)d\eta}\,. \tag{5.1.54}$$

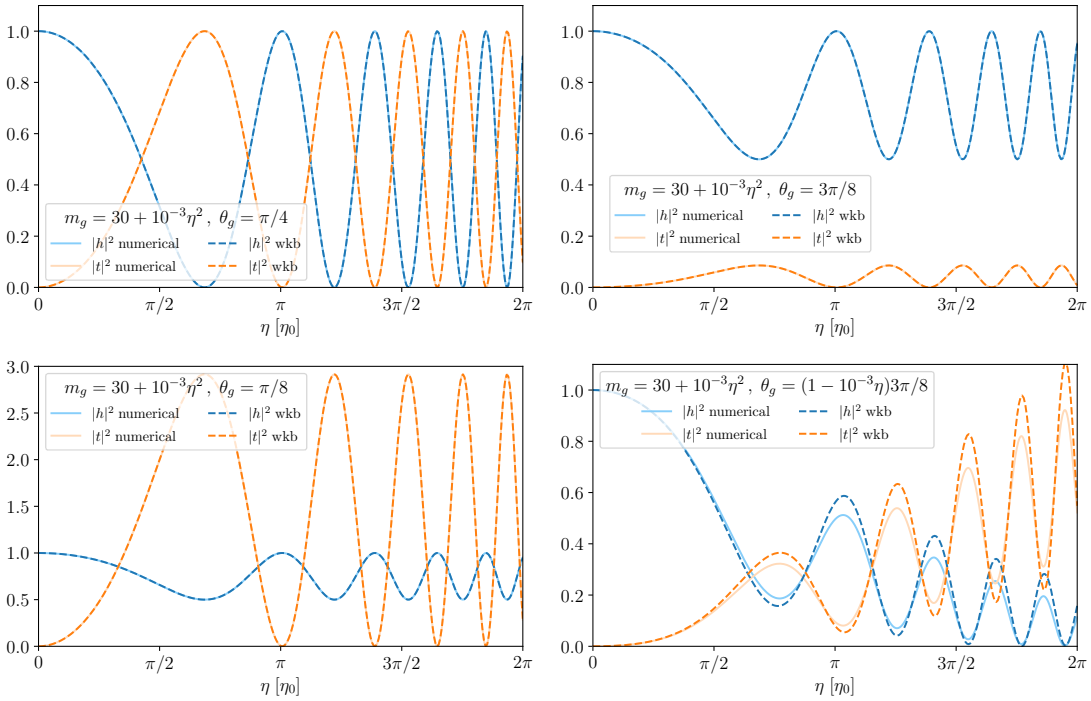

Figure 5.2. Oscillation of the GW amplitude $|h|$ and the tensor perturbation $|t|$ due to a mass mixing. We choose the same time dependent effective mass m_g in all the panels. The mixing angle θ_g corresponds to $\pi/4$ (upper left), $3\pi/8$ (upper right), $\pi/8$ (lower left) and a function of time (lower right). We compare the numerical solution (solid) with the WKB expansion (dashed) fixing $\Delta c = 0$, $k = 10^3$, $h_0 = 1$ and normalizing time w.r.t. the initial period η_0.

This suggests the definition of a mixing angle θ_g,

$$\tan^2\theta_g = -\hat{E}_{12}\hat{E}_{21} = \frac{m_{ht}^2 m_{th}^2}{(m_t^2 + M^2\Delta/2)^2} + \mathcal{O}\left(\frac{\Delta c^2 k^2}{M^2}\right), \tag{5.1.55}$$

so that, after imposing the initial conditions $h(\eta_e) = h_0$ and $t(\eta_e) = 0$, the amplitude of h becomes

$$|h(\eta)|^2 = h_0^2\cos^4\theta_g\left(\frac{\theta_1(\eta_e)}{\theta_1(\eta)} + \frac{\theta_2(\eta_e)}{\theta_2(\eta)}\tan^4\theta_g + 2\frac{\sqrt{\theta_1(\eta_e)\theta_2(\eta_e)}}{\sqrt{\theta_1(\eta)\theta_2(\eta)}}\tan^2\theta_g\cos\left[\int_{\eta_e}^{\eta}\delta\theta(\eta')d\eta'\right]\right), \tag{5.1.56}$$

and the one of t reads

$$|t(\eta)|^2 = h_0^2\cos^4\theta_g\frac{m_{th}^4}{(m_t^2 + M^2\Delta/2)^2}\left(\frac{\theta_1(\eta_e)}{\theta_1(\eta)} + \frac{\theta_2(\eta_e)}{\theta_2(\eta)} - 2\frac{\sqrt{\theta_1(\eta_e)\theta_2(\eta_e)}}{\sqrt{\theta_1(\eta)\theta_2(\eta)}}\cos\left[\int_{\eta_e}^{\eta}\delta\theta(\eta')d\eta'\right]\right). \tag{5.1.57}$$

Therefore, the amplitude of the GW signal $|h|^2$ detected will oscillate with a frequency given by $\delta\theta = m_g^2/(2c_h k)$ and an amplitude controlled by the mixing angle θ_g.

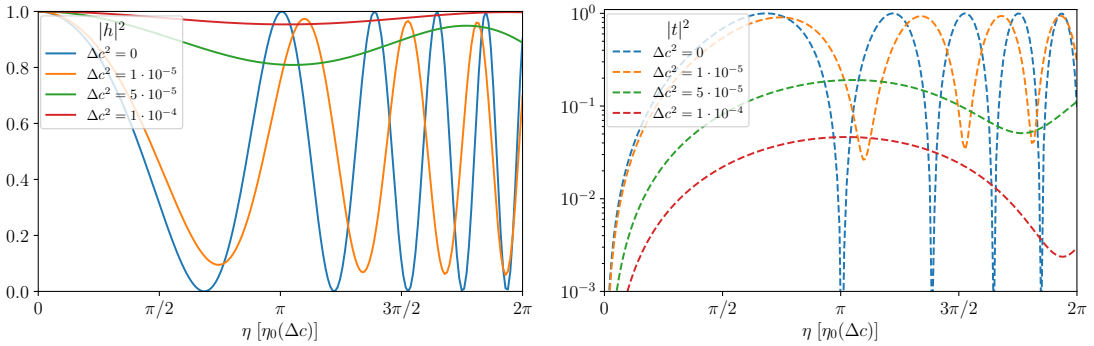

Figure 5.3. Mass mixing with different propagation speeds. We plot the amplitudes $|h|^2$ (left) and $|t|^2$ (right) for different values of $\Delta c^2 = c_t^2 - c_h^2$. We have chosen $\theta_g = \pi/4$, $k = 10^3$, $h_0 = 1$ and the same time dependent effective mass m_g of Fig. 5.2. For each Δc, we have normalized the time w.r.t. its initial period $\eta_0(\Delta c)$.

In Fig. 5.2 we have plotted the evolution of the amplitudes of h and t when there is a mass mixing, in the limit in which the propagation speeds are the same $\Delta c = 0$. We normalize time with respect to the initial period of oscillation $\eta_0 \equiv 1/\delta\theta(\eta_e)$. Although the effective mass m_g varies with time, as can be seen from the change in the frequency of oscillation, the WKB solution (dashed lines) is a very good approximation of the numerical result (solid lines). We notice that the mixing angle determines the amplitude of the second tensor. When $\theta_g = \pi/4$, as in the upper left panel, there is a complete conversion of h into t. As we will see later in section 5.2, this configuration will maximize the detectability of the GW oscillations. When $\theta_g > \pi/4$, as in the upper right panel, there is not a complete conversion and $|t|$ is smaller than h_0. In the opposite case when $\theta_g < \pi/4$, as in the lower left panel, $|t|$ can be larger than h_0. Finally, if θ_g varies in time, as in the lower right panel, the amplitude of $|t|$ will change accordingly in time. In this case, due to the rapid time variation both in $m_g(\eta)$ and $\theta_g(\eta)$ and the choice of $k = 10^3$, the leading order WKB solution does not fully capture the dynamical behavior. For a larger wavenumber, the agreement improves. In this respect, one should remember that for astrophysical sources of GWs and cosmologically varying parameters, the hierarchy in k is many orders of magnitude larger than the one presented in these examples.

In the case in which $\Delta c \neq 0$, there is suppression of the amplitude of t w.r.t. h determined by $m_{ht}^2/(\Delta c^2 k^2)$, recall (5.1.47). This suppression can be observed in Fig. 5.3 where we plot $|h|$ (left) and $|t|$ (right) for different values of Δc. For comparison, we use the same parameters of Fig. 5.2. As the difference in the speeds increases, the amplitude of h approaches the initial value h_0 and the second tensor t reduces. Since we have chosen $k = 10^3$ and $m_{ht} \sim 10$, one needs $\Delta c < 10^{-4}$ not to get a negligible amplitude of t. In practice, for large hierarchies between k and m_{ij}, one needs to have $\Delta c \sim 0$ in order to have observable GW oscillations.

Finally let us remark that not only the amplitude of the GW will differ w.r.t. GR, also the phase will change. From (5.1.48) we learn something important: even if we

set the speed of h equal to the speed of light, $c_h = c$, when there is a mixing and the second tensor propagates at a different speed, the speed of GWs can differ from c. If we define $c_{\rm gw}$ from the leading k^2 term in (5.1.48), we can parametrize the anomalous speed through

$$\alpha_{\rm gw} = \frac{c_{\rm gw}^2}{c^2} - 1 = \frac{1}{2}\frac{\Delta c^2}{c^2}\left(1 - \frac{\Delta m^2}{M^2(1+\Delta)}\right) . \tag{5.1.58}$$

Note that when there is no mixing, m_{th} or m_{ht} vanish, then $M^2(1+\Delta) = \Delta m^2$ and $\alpha_{\rm gw} = 0$. This implies that $\alpha_{\rm gw}$ will be degenerate in Δc and m_{ht}, meaning that $\alpha_{\rm gw} \simeq 0$ whenever $\Delta c \ll 1$ or $m_{ht} \ll 1$.

Mixing through the friction matrix

If the friction is non-diagonal, there will also be GW oscillations. Our starting ansatz is

$$\left[\frac{d^2}{d\eta^2} + \begin{pmatrix} 0 & -2\alpha \\ 2\alpha & 4\Delta\nu \end{pmatrix}\frac{d}{d\eta} + \begin{pmatrix} c_h^2 & 0 \\ 0 & c_t^2 \end{pmatrix} k^2\right]\begin{pmatrix} h \\ t \end{pmatrix} = 0\,, \tag{5.1.59}$$

where α is the parameter controlling the mixing and we have defined $4\Delta\nu = \nu_t - \nu_h$. Note that we make this choice because it is always possible to absorb the part of the friction matrix proportional to the identity via a field redefinition $\vec{\Phi} = e^{-\frac{1}{2}\int \nu_1 d\eta}\tilde{\Phi}$. One might be tempted to proceed similarly and absorb the whole friction matrix with a matrix exponential. However, this is not consistent with the WKB expansion unless this matrix commutes with the matrix of eigenvectors and eigenfrequencies, i.e. $[\hat{E}, \hat{\nu}] = [\hat{\theta}, \hat{\nu}] = 0$.

WKB approximation: Therefore, in general, we will have to solve the quartic equation (5.1.11) for the eigenfrequencies. Although analytically solvable, the solutions themselves are not very illuminating. In the following, for simplicity, we restrict to $\Delta c = 0$.[6] In that case, the eigenfrequencies become

$$\theta_{1,2} = \sqrt{c_h^2 k^2 + (\omega_\nu \pm i\Delta\nu)^2} \pm \omega_\nu + i\Delta\nu\,, \tag{5.1.60}$$

$$\theta_{3,4} = -\sqrt{c_h^2 k^2 + (\omega_\nu \pm i\Delta\nu)^2} \pm \omega_\nu + i\Delta\nu\,, \tag{5.1.61}$$

where we have defined the oscillation frequency $\omega_\nu^2 \equiv \alpha^2 - \Delta\nu^2$ associated to the friction mixing.

Large-k approximation: If we were in a situation in which the parameters are small themselves compared to k, we could apply the large-k approximation. In the $\Delta c = 0$ case, the phases are simply $\theta_{1,2}^2 = c_h^2 k^2$ and $\hat{E} = \hat{I}$. Thus, all the mixing information is contained in the amplitude (5.1.35). Particularizing for the case under consideration, it

[6] In the situation in which $\Delta c \neq 0$, the mixing will be suppressed in analogy to the max mixing case (5.1.47), although with less strength: now $\sim 1/k$ instead of $\sim 1/k^2$.

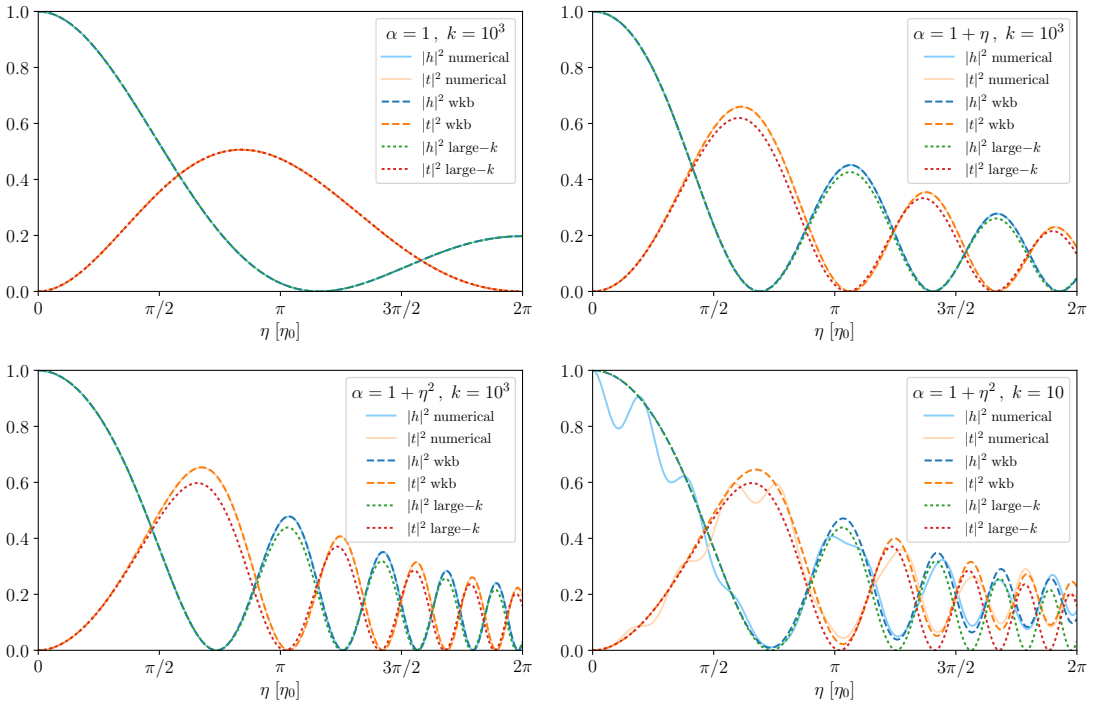

Figure 5.4. Oscillation of the GW amplitude $|h|$ and the tensor perturbation $|t|$ due to a friction mixing. We vary the time dependence of the non-diagonal matrix element α from constant (upper left), to linear (upper right), to quadratic (lower panels). We compare the numerical solution (solid) with the WKB expansion (dashed) and the large-k approximation (dotted) fixing $\Delta c = 0$, $h_0 = 1$ and $k = 10^3$, except in the lower left plot where $k = 10$. We normalize time w.r.t. the initial period η_0.

becomes

$$\begin{aligned}\vec{\Phi}_0 &= \frac{\sqrt{c_h(\eta_e)}}{\sqrt{c_h(\eta)}} e^{-\frac{1}{2}\int \hat{\nu} d\eta} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\ &= \frac{\sqrt{c_h(\eta_e)}}{\sqrt{c_h(\eta)}} e^{-\bar{\Delta\nu}} \begin{pmatrix} \cos\bar{\omega}_\nu + \frac{\bar{\Delta\nu}}{\bar{\omega}_\nu}\sin\bar{\omega}_\nu & \frac{\bar{\alpha}}{\bar{\omega}_\nu}\sin\bar{\omega}_\nu \\ -\frac{\bar{\alpha}}{\bar{\omega}_\nu}\sin\bar{\omega}_\nu & \cos\bar{\omega}_\nu - \frac{\bar{\Delta\nu}}{\bar{\omega}_\nu}\sin\bar{\omega}_\nu \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix},\end{aligned} \tag{5.1.62}$$

where we have defined the integrals

$$\bar{\Delta\nu} = \int_{\eta_e}^{\eta} \Delta\nu d\eta\,, \qquad \bar{\alpha} = \int_{\eta_e}^{\eta} \alpha d\eta\,, \qquad \bar{\omega}_\nu^2 = \bar{\alpha}^2 - \bar{\Delta\nu}^2\,, \tag{5.1.63}$$

from the time of emission η_e to a given instant η. Then, imposing the initial conditions, $h(\eta_e) = h_0$ and $t(\eta_e) = 0$, we get

$$|h(\eta)|^2 = h_0^2 \frac{c_h(\eta_e)}{c_h(\eta)} e^{-2\bar{\Delta\nu}} \left(\cos\bar{\omega}_\nu + \frac{\bar{\Delta\nu}}{\bar{\omega}_\nu} \sin\bar{\omega}_\nu \right)^2 , \tag{5.1.64}$$

$$|t(\eta)|^2 = h_0^2 \frac{c_h(\eta_e)}{c_h(\eta)} e^{-2\bar{\Delta\nu}} \frac{\bar{\alpha}^2}{\bar{\omega}_\nu^2} \sin^2\bar{\omega}_\nu \, . \tag{5.1.65}$$

Therefore, the frequency of oscillation is controlled by $\bar{\omega}_\nu$ and the damping of the signal by $\bar{\Delta\nu}$.

In the simplest case in which the friction matrix $\hat{\nu}$ is constant, $\alpha = \alpha_0$ and $\Delta\nu = \Delta\nu_0$, the integrals (5.1.63) simplify to $\bar{\Delta\nu} = \Delta\nu_0(\eta - \eta_e)$, $\bar{\alpha} = \alpha_0(\eta - \eta_e)$ and $\bar{\omega}_\nu = \sqrt{\alpha_0^2 - \Delta\nu_0^2}(\eta - \eta_e) = \omega_0(\eta - \eta_e)$. Accordingly, only the terms in the sines and cosines, and the global damping depend on time

$$|h(\eta)|^2 = h_0^2 \frac{c_h(\eta_e)}{c_h(\eta)} e^{-2\Delta\nu_0\bar{\eta}} \left(\cos\left[\omega_0\bar{\eta}\right] + \frac{\Delta\nu_0}{\omega_0} \sin\left[\omega_0\bar{\eta}\right] \right)^2 , \tag{5.1.66}$$

$$|t(\eta)|^2 = h_0^2 \frac{c_h(\eta_e)}{c_h(\eta)} e^{-2\Delta\nu_0\bar{\eta}} \frac{\alpha_0^2}{\omega_0^2} \sin^2\left[\omega_0\bar{\eta}\right] , \tag{5.1.67}$$

where $\bar{\eta} = (\eta - \eta_e)$. This solution resembles the model of GW-gauge field oscillations studied in Ref. [319, 396].

In Fig. 5.4 we plot different examples of the oscillations and damping in the amplitude of the tensor perturbations h and t induced by the friction mixing. In order to compare the numerical solution (solid lines) with the WKB expansion (dashed lines) and the large-k approximation (dotted lines), we consider different time dependences of the mixing parameter α and different wave-numbers k. For $k = 10^3$, we observe that the leading WKB solution gives and excellent approximation of the numerical result for a constant, linear and quadratic dependence in time of the mixing α (upper left, upper right and lower left panels respectively). On the contrary, for this wavenumber, the large-k expansion does not match perfectly the numerical result when there is a time dependence. In order to find departures of the WKB and the numerical solution, we have to lower the value of the wavenumber to $k = 10$ (lower right panel). This serves to exemplify that the WKB is a better approximation in general than the large-k expansion since it expands over the variation of the parameters w.r.t. the frequency and not the parameters themselves. However, when there is a large value of k, both approximations tends to converge and the large-k expansion becomes more useful since the analytical expressions are simpler.

Mixing through the velocity matrix

Certain operators introduce also a non-diagonal velocity matrix. In that case, a mixing occurs at leading order in both WKB and large-k expansions. Focusing only in this source of mixing,

$$\left[\frac{d^2}{d\eta^2} + \begin{pmatrix} c_h^2 & c_{ht}^2 \\ c_{ht}^2 & c_t^2 \end{pmatrix} k^2 \right] \begin{pmatrix} h \\ t \end{pmatrix} = 0 \, , \tag{5.1.68}$$

we can easily solve the propagation. Using the previous convention $\Delta c^2 = c_t^2 - c_h^2$, the eigenfrequencies are given by

$$\theta_{1,2}^2 = \left(c_h^2 + \frac{1}{2}\Delta c^2 \mp \frac{1}{2}\sqrt{4c_{ht}^4 + \Delta c^4} \right) k^2 \,, \tag{5.1.69}$$

and the matrix of eigenvectors by

$$\hat{E} = \begin{pmatrix} 1 & -\frac{2c_{ht}^2}{\Delta c^2 + \sqrt{4c_{ht}^4 + \Delta c^4}} \\ \frac{2c_{ht}^2}{\Delta c^2 + \sqrt{4c_{ht}^4 + \Delta c^4}} & 1 \end{pmatrix} . \tag{5.1.70}$$

Then, the the mixing in the amplitude is determined by the non-diagonal entry c_{ht}. Note here that although a different propagation speed $\Delta c \neq 0$ tend to suppress the mixing, this is not enhanced by k as in previous cases, for instance for the mass mixing case (5.1.47), the non-diagonal terms were suppressed by $\sim 1/k^2$. This can be seen in Fig. 5.5, where we present the oscillation in h and t for different values of Δc in the left and right plot respectively. The larger Δc becomes, the more $|t|$ is suppressed and the more $|h|$ approaches the initial value h_0. This plot is analogous to the mass mixing case presented in Fig. 5.3. However, for the velocity mixing the suppression is not enhanced by k and, thus, Δc can be larger. This different behavior becomes more extreme as k grows.

As in the previous cases, even if $c_h = c$, there can be an anomalous speed $c_{\rm gw} \neq c$ whenever there is a mixing via $c_{ht} \neq 0$ and the second tensor t has a non-luminal propagation speed $c_t \neq c$, as it can be easily deduced from

$$\alpha_{\rm gw} = \frac{1}{2}\frac{\Delta c^2}{c^2}\left(1 - \sqrt{1 + 4\frac{c_{ht}^4}{\Delta c^4}} \right) . \tag{5.1.71}$$

The anomalous speed $\alpha_{\rm gw}$ is thus degenerate in the difference of the speeds Δc and the mixing term c_{ht}.

Chiral mixing

Whenever there is a parity violating term, it is convenient to work in the left- and right-circular polarizations, which we assume in the following. In the simplest set-up, there is only the parity violating matrix $\hat{\Pi}$ linear in k and the velocity matrix,

$$\left[\frac{d^2}{d\eta^2} + \begin{pmatrix} c_h^2 & 0 \\ 0 & c_t^2 \end{pmatrix} k^2 \pm \begin{pmatrix} \mu_h & \gamma \\ \gamma & \mu_t \end{pmatrix} k \right] \begin{pmatrix} h_{L,R} \\ t_{L,R} \end{pmatrix} = 0 \,. \tag{5.1.72}$$

Due to the $\pm$ in front of the $\hat{\Pi}$ matrix, the L and R polarizations evolve differently. As in the previous cases, we compare different approximate solutions of these coupled differential equations. We will see that the WKB analysis will be similar to the mass mixing studied in section 5.1.2 and the large-k similar to the friction mixing studied in section 5.1.2.

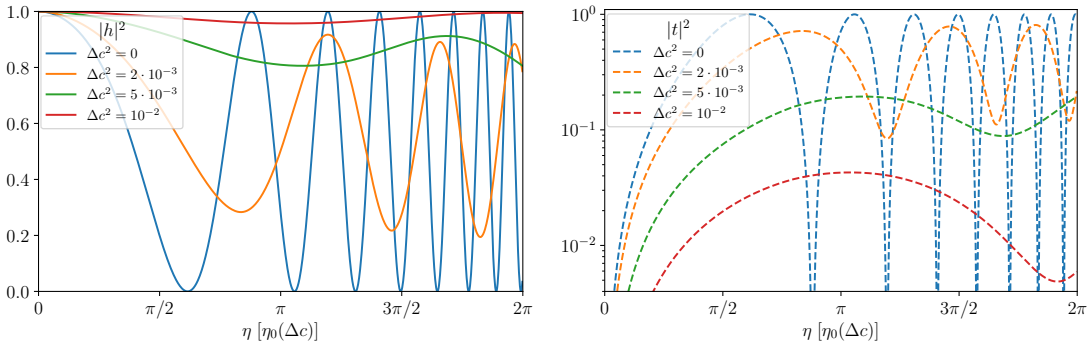

Figure 5.5. Oscillation of the GW amplitude $|h|$ (left) and the tensor perturbation $|t|$ (right) due to a velocity mixing for different values of $\Delta c^2 = c_t^2 - c_h^2$. We have chosen $k = 10^5$, $h_0 = 1$ and the mixing c_{ht} quadratic in time. For each Δc, we have normalized the time w.r.t. its initial period $\eta_0(\Delta c)$.

WKB approximation: at leading order in the WKB we obtain the phase of the wave by solving the algebraic equation (5.1.11). The corresponding phases for each polarization are

$$(\theta_{L,R;\,1})^2 = c_h^2 k^2 + \frac{\Delta c^2 k^2}{2} \pm \frac{\mu_{\rm tot} k}{2} - \frac{k}{2}\sqrt{4\gamma^2 + (\Delta\mu \mp k\Delta c^2)^2}\,, \tag{5.1.73}$$

$$(\theta_{L,R;\,2})^2 = c_h^2 k^2 + \frac{\Delta c^2 k^2}{2} \pm \frac{\mu_{\rm tot} k}{2} + \frac{k}{2}\sqrt{4\gamma^2 + (\Delta\mu \mp k\Delta c^2)^2}\,, \tag{5.1.74}$$

where we have defined $\mu_{\rm tot} = \mu_h + \mu_t$ and $\Delta\mu = \mu_t - \mu_h$. The matrix of eigenvectors is analogous to the mass mixing case (5.1.44) substituing $\hat{M}$ for $\hat{\Pi}\, k$ and accounting for the different sign of the parameters of each polarization

$$\hat{E}_{L,R} = \begin{pmatrix} 1 & \mp\frac{\gamma}{c_h^2 k^2 \pm \mu_h - \theta_2^2} \\ \mp\frac{\gamma}{c_t^2 k^2 \pm \mu_t - \theta_1^2} & 1 \end{pmatrix}. \tag{5.1.75}$$

With these expressions one can proceed and analyze how a GW signal will be modified. There will be both a modification of the amplitude and the phase due to the GW oscillations. These modifications will depend on the polarization. In fact we can already anticipate from the matrix (5.1.75) that there will be a chiral effect in the amplitude. We will discuss on section 5.2.4 how to probe this chirality. For the moment, let us focus on the phase. In the limit in which the difference in the propagation speeds Δc is small, we can extract the GW speed from the leading k^2 term. Although h propagates at the speed of light, the non-luminal speed of t together with the mixing γ induces an anomalous speed for the GWs, parametrized by $\alpha_{\rm gw} = c_{\rm gw}^2/c^2 - 1$. We obtain

$$(\alpha_{\rm gw})_{L,R} = \frac{1}{2}\frac{\Delta c^2}{c^2}\left(1 \mp \frac{\Delta\mu}{4\omega_\mu}\right). \tag{5.1.76}$$

where, for later convenience, we have introduced the frequency $16\omega_\mu^2 = 4\gamma^2 + \Delta\mu^2$.

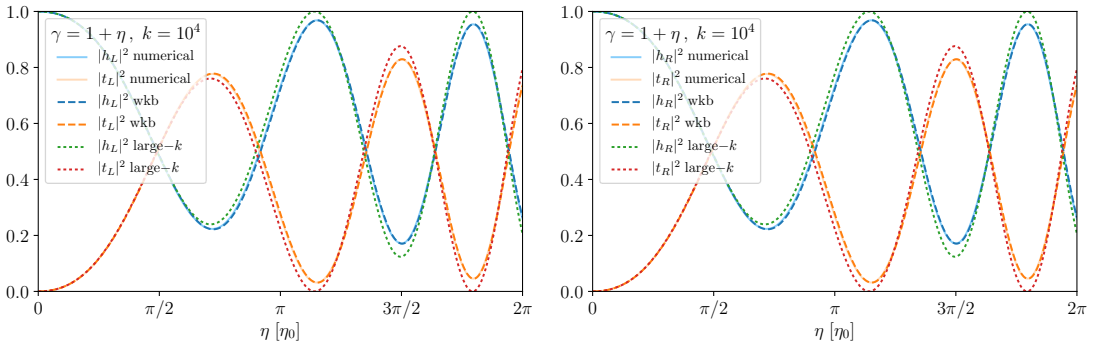

Figure 5.6. Oscillation of the circular polarizations (left and right) of the GW amplitude $|h_{L,R}|$ and the tensor perturbation $|t_{L,R}|$ due to a chiral mixing. We choose the mixing γ to vary linearly in time. We compare the numerical solution (solid) with the WKB expansion (dashed) and the large-k approximation (dotted) fixing $\Delta c = 0$, $h_0 = 1$, $k = 10^4$ and normalizing time w.r.t. the initial period η_0.

Large-k approximation: let us now consider the limit in which both perturbations propagate at the same speed, $\Delta c = 0$. Then, similarly to the friction mixing, in the large-k expansion one obtains $\theta^2_{1,2} = c_h^2 k^2$ and $\hat{E} = \hat{I}$ for both polarizations. Using (5.1.30), we obtain the leading order amplitude $\vec{\Phi}_0$. One may notice that the situation is equivalent to the friction mixing if we exchange $\hat{\nu} \to -i\, k\hat{\Pi}\hat{\theta}^{-1}$. Thus we get

$$\begin{aligned}\left(\vec{\Phi}_0\right)_{L,R} &= \frac{\sqrt{c_h(\eta_e)}}{\sqrt{c_h(\eta)}} e^{\pm\frac{i}{2}\int \hat{\Pi}\, c_h^{-1} d\eta} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\ &= \frac{\sqrt{c_h(\eta_e)}}{\sqrt{c_h(\eta)}} e^{\pm\frac{i}{4}\bar{\mu}_{\rm tot}} \begin{pmatrix} \cos\bar{\omega}_\mu \mp i\frac{\bar{\Delta\mu}}{4\bar{\omega}_\mu}\sin\bar{\omega}_\mu & \pm i\frac{\bar{\gamma}}{2\bar{\omega}_\mu}\sin\bar{\omega}_\mu \\ \pm i\frac{\bar{\gamma}}{2\bar{\omega}_\mu}\sin\bar{\omega}_\mu & \cos\bar{\omega}_\mu \pm i\frac{\bar{\Delta\mu}}{4\bar{\omega}_\mu}\sin\bar{\omega}_\mu \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix},\end{aligned} \tag{5.1.77}$$

where we have defined the integrals

$$\bar{\mu}_{\rm tot} = \int_{\eta_e}^{\eta} \frac{\mu_{\rm tot}}{c_h} d\eta\,, \quad \bar{\Delta\mu} = \int_{\eta_e}^{\eta} \frac{\Delta\mu}{c_h} d\eta\,, \quad \bar{\gamma} = \int_{\eta_e}^{\eta} \frac{\gamma}{c_h} d\eta\,, \quad 16\bar{\omega}_\mu^2 = 4\bar{\gamma}^2 + \bar{\Delta\mu}^2\,. \tag{5.1.78}$$

Imposing the initial conditions, $h(\eta_e) = h_0$ and $t(\eta_e) = 0$, we obtain the amplitude of the tensor perturbations

$$|h_{L,R}|^2 = h_0^2 \frac{c_h(\eta_e)}{c_h(\eta)} \left(\cos^2\bar{\omega}_\mu + \frac{\bar{\Delta\mu}^2}{16\bar{\omega}_\mu^2}\sin^2\bar{\omega}_\mu\right), \tag{5.1.79}$$

$$|t_{L,R}|^2 = h_0^2 \frac{c_h(\eta_e)}{c_h(\eta)} \frac{\bar{\gamma}^2}{4\bar{\omega}_\mu^2}\sin^2\bar{\omega}_\mu\,. \tag{5.1.80}$$

Noticeably, the amplitude is the same for both polarizations. This means that in this setup, at leading order in the large-k expansion, there is no chiral effect. Note however

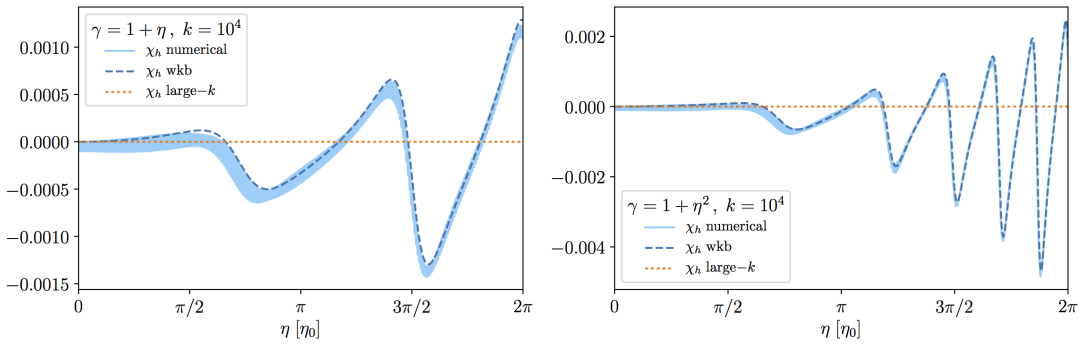

Figure 5.7. Chirality induced by GW oscillations. We compare the value of χ_h, see (5.1.81), for the numerical (solid), WKB (dashed) and large-k (dotted) solutions. We plot a linear and a quadratic time dependent mixing γ in the left and right panels respectively, fixing $\Delta c = 0$, $h_0 = 1$, $k = 10^4$ as in Fig. 5.6.

that if instead of starting only with a chiral matrix $\hat{\Pi}$, we include a friction matrix $\hat{\nu}$, there will be chiral effects. This is because the matrix in the exponent of (5.1.30) will then be $\hat{A}_{\text{large-k}} = \hat{\nu} - i\,k\hat{\Pi}\hat{\theta}^{-1}$. Another point to highlight is that, differently to the friction mixing, now there is not a global damping because the $e^{\pm\frac{i}{4}\bar{\mu}_{\text{tot}}}$ term only contributes to the phase.

In Fig. 5.6 we plot the amplitude of the two polarizations (left and right panels respectively) of the perturbations of h and t. We choose a mixing parameter that varies linearly in time. As for the friction mixing, the WKB solution (dashed lines) provides a better approximation of the numerical result (solid lines) than the large-k expansion (dotted lines). Moreover we observe that both the left and right polarizations evolve qualitatively in the same manner. To quantify the difference in the evolution, we introduce the chirality parameter χ. For the GW amplitude, we define it as

$$\chi_h = \frac{|h_L|^2 - |h_R|^2}{|h_L|^2 + |h_R|^2}\,. \tag{5.1.81}$$

In Fig. 5.7 we plot χ_h for the numerical, WKB and large-k solutions. As discussed, at leading order in the large-k expansion, there is no chirality. On the other hand, the WKB gives a good agreement with the numerical result. An important point clearly seen in the plots is that the chirality is an accumulative effect, thus growing along the propagation. This means that even if the parameters are small, the effect can become dominant over long travel distances. We compare the growth of the chirality for a linear (left panel) and quadratic (right panel) time dependence of the chiral mixing γ. We will study the detectability of this effect in section 5.2.4.

5.2 Phenomenology and constraints

So far we have shown that the coupled evolution of two tensor perturbations (5.1.2) can modify their propagation in several different ways:

(i) their amplitudes mix whenever there are non-diagonal terms in the mass, friction, velocity or chiral matrices;

(ii) their amplitude can get damped due to the friction matrix;

(iii) the propagation speed of the perturbation coupled to matter could be anomalous whenever the second perturbation has a non-luminal propagation and there is a mixing;

(iv) each polarization propagates differently when there is a chiral matrix.

This type of propagation equations arises in cosmological set-ups with multiple vector fields and in many classes of modified gravity theories. For instance, bigravity leads to a mass mixing while cosmological gauge fields yields to a friction and chiral mixing. More sophisticated vector-tensor theories can even induce a velocity mixing.

In this section we are going to investigate the phenomenological implications of these effects on GW observations. In order to discuss each possible observable, we will consider representative examples. Here, we do not aim at setting firm constraints on particular theories but rather show the potential of GW oscillations to test certain classes of models. A detailed analysis solving the full cosmological evolution together with GW propagations for each particular example would be necessary for that task and is left for the interested readers to test their favorite theory.

The mixing of the amplitude of the different tensor perturbations has clear consequences for the GW signals. Even if we start only with perturbations of one class at emission, we will generically have both of them excited at detection. Since only one of the perturbations, h in our convention, couples to matter, the excitation of the other perturbation t would be seen in the detector as a deficit of h signal. If the conversion of h into t continues periodically, this will induce an oscillation of the GW wave-form. We will study this characteristic effect in section 5.2.1. But, if the amplitude detected is lower, this would be interpreted as the source being further away. Therefore, there will be also a modification of the GW luminosity distance that we analyze in section 5.2.2. Now, if t propagates at a speed different from the speed of light c, the mixing produces that the net propagation speed of h becomes anomalous. This can be constrained with multi-messenger detections as we discuss in section 5.2.3. Finally, for the case in which the mixing is chiral, the two tensor polarizations $h_{+,\times}$ evolve differently leaving an imprint that could be distinguished with a network of ground based detectors. We study this imprint in section 5.2.4.

5.2.1 Oscillations of the wave-form

The mixing of the tensor perturbations produces that the GW strain of the signal emitted is modified during the propagation. This modification will depend on the particular theory and on the location of the source. To exemplify this effect, we are going to consider two representative examples.

On the one hand, we are going to investigate a scenario in which there is a *mass mixing* whose time dependence is proportional to the square of the scale factor, i.e.

$\hat{M} \propto a(\eta)^2 \hat{M}_0$. We will work in the high-k limit in which $\theta_1 \simeq \theta_2 \simeq k$. Accordingly, we parametrize the problem with an effective mass m_g and a mixing angle θ_g which are constants, recall (5.1.52) and (5.1.55) respectively. This example resembles bigravity theory in the large mass limit, which will be presented in more detail in section 5.3. The transfer function between the initial GR emission $|h_{GR}|$ and the signal detected $|h(z)|$ is given by

$$\frac{|h(z,k)|}{|h_{GR}|} = \cos^2\theta_g \left(1 + \tan^4\theta_g + 2\tan^2\theta_g \cos\left[\frac{m_g^2}{2k}\int_0^z \frac{dz}{(1+z)^2 H(z)}\right]\right)^{1/2} . \quad (5.2.1)$$

Importantly, the transfer function depends in the parameters of the model, m_g and θ_g, the redshift z and the frequency k. In addition, the modified amplitude is also sensitive to the cosmic expansion history through $H(z)$. As observations suggest and for simplicity, we have assumed that the background cosmology is ΛCDM.

On the other hand, we will work with an example with a *friction mixing*. In this case, we consider that the friction matrix $\hat{\nu}$ is constant, which corresponds to (5.1.66). The transfer function

$$\frac{|h(z)|}{|h_{GR}|} = (1+z)^{-\Delta\nu}\left(\cos\left[\omega_\nu \log(1+z)\right] + \frac{\Delta\nu}{\omega_\nu}\sin\left[\omega_\nu \log(1+z)\right]\right) \quad (5.2.2)$$

is then controlled by the frequency of oscillation ω_ν and the damping factor $\Delta\nu$, where ω_ν is a function of both the non-diagonal entry α and $\Delta\nu$. Noticeably, now the amplitude does not depend on the frequency of the GW.

In Fig. 5.8 we compare the modification of the GW strain for the mass and friction mixing scenarios. We plot the modified strain on top of the original GR signal both in the frequency and time domain. As it can be clearly seen, the fact that the mass mixing transfer function depends on the frequency makes the strain to oscillate leaving a very distinct wave-form. On the contrary, for the friction mixing there is only a dimming of the signal that globally rescales the amplitude. This effect is completely degenerate with the distance to the source and could not be distinguished through wave-form modeling. Still, since the transfer function depends on redshift, a friction mixing leaves a measurable imprint in the GW luminosity distance as we will see in section 5.2.2.

Focusing on the mass mixing case, the oscillatory pattern in the GW strain could be used to constrain the parameters of the theory by comparing this wave-form with the observed ones. In this sense, the best target will be a signal with a long inspiral part, allowing to constrain the strain over many oscillations. In the context of present ground-based detectors, long binary neutron star signals like GW170817 have more constraining power that short binary black-hole detections like GW150914. Moreover, with the future space-based detector LISA, we could be sensitive to much lower frequencies of oscillations. We will also benefit from very long signals that could last months or years. Eventually, a multi-band GW detection could be as well a very powerful test of this type of mixing.

Lastly, let us emphasis the importance of searching for possible astrophysical degeneracies that could mimic this fundamental oscillatory pattern. In particular, we note

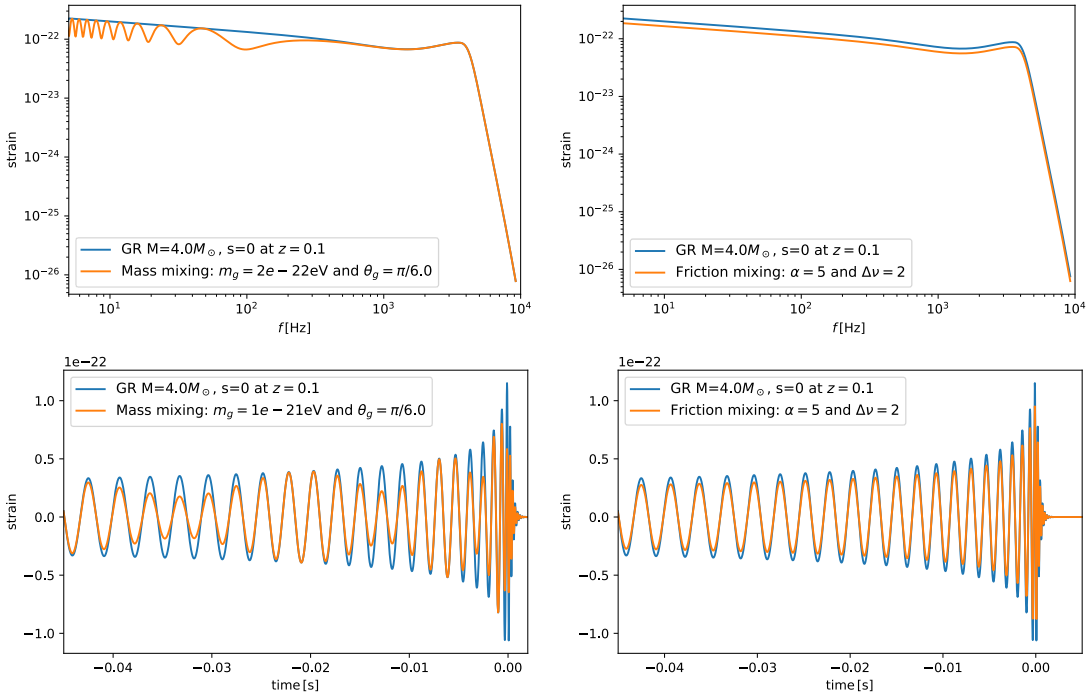

Figure 5.8. Examples of modified wave-forms in the frequency (top) and time (bottom) domain for theories with a mass mixing (left) or a friction mixing (right). The emitted GR signal (blue) corresponds to an equal mass, non-spinning BBH merger of $M = 4M_\odot$ at $z = 0.1$.

that binaries with a precessing spins also lead to an oscillation of the wave-form. The oscillation effect is enhanced when there is a hierarchy in the masses of the compact objects. Thus, this will be more relevant to LISA sources.

5.2.2 Modified GW luminosity distance

If part of the initial GW signal is converted into the second tensor or diluted by the friction term, this would be interpreted as the source being further apart since the received signal would be dimmer. In other words, the GW luminosity distance $d_L^{\rm gw}$ would be modified. In chapter 4 we studied the modification of $d_L^{\rm gw}$ in gravity theories with a single tensor perturbations. Here we will extend this analysis to allow for GW oscillations.

We saw in section 4.2 that there are two ways in which we can test $d_L^{\rm gw}$: with GWs alone and with multi-messenger events. In the first case, one could use GW number counts as a function of redshift. However, for a monotonic behavior of $d_L^{\rm gw}$ it might be difficult to disentangle this effect from the uncertainty in the astrophysical modeling of the merger rate. On the contrary, when there is a mixing of the perturbations, for given redshifts periodically separated, the number of detections will be much smaller than predicted. This distinct pattern could be very well distinguished from an astrophysical

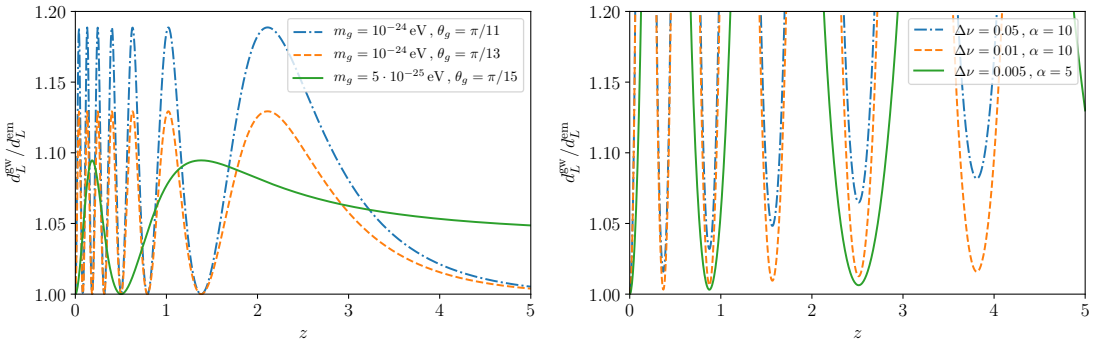

Figure 5.9. Modified GW luminosity distance as a function of redshift for theories with a mass mixing (left) or a friction mixing (right). We plot the ratio $d_L^{\mathrm{gw}}/d_L^{\mathrm{em}}$ for different values of the effective mass m_g and mixing angle θ_g; and damping factor $\Delta\nu$ and friction mixing α.

effect. In the extreme case in which there is a complete conversion of h into t (see for instance Fig. 5.2), this implies that $R(z)$ at certain redshift bins would be zero. Having a large population of compact binaries over a wide range of redshift could bound the mixing of the GWs with other tensor modes.

With multi-messenger events, we have more constraining power. The advantage is that we can probe the ratio $d_L^{\mathrm{gw}}/d_L^{\mathrm{em}}$. Any deviation of this ratio from being 1 would be a sign of physics beyond GR and ΛCDM. Directly from the transfer function in the GW amplitude we can compute the ratio of luminosity distances. For the mass mixing described in the previous section it would be given by

$$\frac{d_L^{\mathrm{gw}}}{d_L^{\mathrm{em}}} = \frac{1}{\cos^2\theta_g}\left(1+\tan^4\theta_g + 2\tan^2\theta_g\,\cos\left[\frac{m_g^2}{2k}\int_0^z \frac{dz}{(1+z)^2 H(z)}\right]\right)^{-1/2}. \tag{5.2.3}$$

Possible mixing angles span form 0 to $\pi/2$, having the maximum mixing at $\theta_g = \pi/4$. At this value, a complete conversion of h into t can occur. As a consequence, the amplitude of h vanishes and d_L^{gw} diverges at

$$\frac{m_g^2}{2k}\int_0^z \frac{dz}{(1+z)^2 H(z)} = \frac{\pi}{2} + 2\pi n \quad \text{for} \quad n = 0, 1, 2, \cdots \tag{5.2.4}$$

Note also that the parameter space is symmetric around $\pi/4$. The frequency of oscillation is controlled by the effective mass and the frequency of the GW through m_g^2/k. To have a frequency of oscillation of order 1 at low redshift, this ratio should be of order H_0 to compensate the Hubble parameter in the denominator.[7] This implies that with present ground-based interferometer, $f_{\mathrm{LIGO}} \sim 100\mathrm{Hz}$, we can test $m_g \sim 10^{-23}\mathrm{eV}$. In the same manner, the future space-based detector LISA, $f_{\mathrm{LISA}} \sim 10\mathrm{mHz}$, will be sensitive to $m_g \sim 10^{-25}\mathrm{eV}$.

[7]The reader should remember that $H_0 \sim 10^{-33}\mathrm{eV} \sim 10^{-18}\mathrm{Hz}$.

For the friction mixing example, the ratio of the GW and EM luminosity distance becomes

$$\frac{d_L^{\rm gw}}{d_L^{\rm em}} = (1+z)^{\Delta\nu}\left(\cos\left[\omega_\nu\,\log(1+z)\right] + \frac{\Delta\nu}{\omega_\nu}\sin\left[\omega_\nu\,\log(1+z)\right]\right)^{-1}\,. \tag{5.2.5}$$

The first term produces a global friction term that increases $d_L^{\rm gw}$, while the rest makes the luminosity distance to oscillate at a rate determined by $\omega_\nu = \sqrt{\alpha^2 - \Delta\nu^2}$, where α is the non-diagonal term producing the mixing and $\Delta\nu$ is the difference between the friction term of t and h. Note that since we are taking the ratio of $d_L^{\rm gw}$ over $d_L^{\rm em}$, the standard damping due to the cosmic expansion does not appear in this expression. Moreover, in the limiting case in which there is no mixing, $\alpha \to 0$, we recover $d_L^{\rm gw}/d_L^{\rm em} \to 1$. Contrary to the mass mixing case, here there is always a complete conversion of one type of perturbation into the oder happening at

$$\tan\left[\omega_\nu\,\log(1+z)\right] = -\frac{\omega_\nu}{\Delta\nu}\,. \tag{5.2.6}$$

This extreme feature in $d_L^{\rm gw}/d_L^{\rm em}$ makes it easier to probe.

In Fig. 5.9 we have plotted the ratio $d_L^{\rm gw}/d_L^{\rm em}$ for both the mass mixing and friction mixing scenarios as a function of redshift. In the left panel one can observe how the GW luminosity distance varies with θ_g and m_g. As the mixing angle approaches to $\pi/4$, the amplitude of the oscillation increases. Accordingly, the frequency of oscillation increases with m_g. On the right panel we present the corresponding plot for the friction mixing. Noticeably, the ratio diverges periodically due to the complete conversion of the original signal into tensor perturbations not coupled to matter. Also, the global friction term $(1+z)^{\Delta\nu}$ makes the minimum values of $d_L^{\rm gw}/d_L^{\rm em}$ to increase away from 1 as $\Delta\nu$ increase. On the other hand, the mixing parameter α controls the frequency of oscillation.

Detecting standard sirens over different redshift ranges allows to cover a larger patch of the parameter space. Second generation ground-based interferometers are sensitive to BNS up to $z \sim 0.05$ and BBH up to $z \sim 0.5$. This range will be much increased with third generation detectors such as Einstein Telescope, reaching possibly $z \sim 2$ for BNS and $z \sim 15$ for BBH. LISA from space could also hear up to very high redshift. Fig. 5.9 corresponds to the expected redshift range, $z \sim 2-6$, and sensitivity, $\Delta d_L/d_L \sim 10\%$, where LISA could detect standard sirens from super massive BHs with EM counterparts. We will study in more detail in section 5.3 the capability of LISA to detect modifications in $d_L^{\rm gw}$, showing that for the case of bigravity LISA could probe masses of $m_g \gtrsim 2{\cdot}10^{-25}$eV with mixing angles $0.05\pi \lesssim \theta_g \lesssim 0.45\pi$. In the same manner, we note that LISA could also probe scenarios with a friction mixing.

5.2.3 Anomalous GW speed

For the moment, we have focused on modifications of the amplitude of GWs due to the mixing of the tensor modes coupled to matter with other cosmological tensor fields. Nonetheless, GW oscillations can also modified the phase of the GW, which is indeed much better constrained with interferometers than the amplitude. One of the results

of our analysis was to demonstrate that even if the tensor mode coupled to matter h propagates at the speed of light $c_h = c$, if there is a mixing and the second tensor propagates at a different speed, the effective velocity of the GW could be non-luminal, $c_{\rm gw} \neq c$. We have shown this explicitly for the mass-mixing (5.1.58) and velocity mixing (5.1.71) scenarios.

As we discussed in detail in chapter 3, an anomalous speed, which can be parametrized by $\alpha_{\rm gw} = c_{\rm gw}^2/c^2 - 1$, yields to a delay between the GW and any other EM counterpart Δt. After GW170817, we know that $\alpha_{\rm gw}$ is constrained to the level of 10^{-15} at LIGO frequencies. Assuming that h propagates at the speed of light, this multi-messenger event constrains the possible deviation from c of the second tensor whenever there is mixing. Suppressing the mixing could also be a way to avoid this limit but then there will be no other GW oscillation effects in the amplitude.

Present constraints on the propagation speed of GWs could be improved in the future by observing more distant events (remember that GW170817 was only at about 40Mpc). The increase in the sensitivity and distance reach will be more significant when moving from second to third generation interferometers. Also promising candidates are the SMBH standard siren at high redshift that LISA target to detect, although it is still not clear if prompt emission could be observed for such distant sources. Interestingly, LISA also provides the opportunity to test the propagation speed of GWs at a different frequency range, which is relevant to constrain a possible frequency dependence in $c_{\rm gw}$ [7, 389]. A secure test of the speed of GWs at mHz is to measure the phase lag between GW and EM radiation of LISA verification binaries [2, 333].

5.2.4 Chirality

Finally, let us examine how to probe modifications in the propagation of the different polarizations $h_{+,\times}$ due to chiral GW oscillations. For that, one needs to be sensitive to each polarization. However, in general, the polarizations are degenerate with the location in the sky and the inclination angle (see for instance Eq. (2.2.8)). Having a network of ground-based detectors across the surface of the Earth can break this degeneracy. Moreover, if the source is located with an EM counterpart, the capability to probe different polarizations increases. Since the two LIGO detectors are aligned to maximized the joint sensitivity, the role of Virgo has been crucial to start performing tests of the types of polarizations. For instance, with the three-detector detection GW170814 [348] it was possible to contrast the hypothesis of the signal being purely tensor against it being purely vector or scalar. A much stronger result favoring purely tensor polarizations was obtained with GW170817 [391] since the sky position was determined with high accuracy thanks to the EM counterparts.

In order to test chiral GW oscillations it would be needed to distinguish each polarization and measure their amplitude as a function of frequency $h_{+,\times}(f)$. Alternatively, one could measure the luminosity distance as a function of redshift for each polarization. In principle, with a network of detectors these effects could be probed. However, one should remember that this is going to be a small effect since it is suppressed by the wavenumber. In fact, in the large-k (or shortwave) limit, we have seen that the chirality

χ_h vanishes (see discussion in section 5.1.2). We leave the study of particular scenarios in which the chiral mixing is enhanced for future work.

5.3 Probing bigravity with LISA standard sirens

After presenting a general formalism to study GW oscillations and investigating the possible observational effects, let us now concentrate on a representative case: bigravity theory. A consistent theory of bigravity, free of Ostrogradsky instabilities at the fully non-linear level, has been proposed by Hassan and Rosen (HR) [150], adding an Einstein-Hilbert term for the reference metric of the so-called *dRGT* theory of massive gravity [137]. Bigravity admits FRW cosmological solutions describing late-time acceleration that differ from ΛCDM at the level of the background as well as for the dynamics of cosmological fluctuations. In this section we study the evolution of tensor modes around homogeneous FRW configurations, applying the techniques developed in section 5.1.1 that will allow us to go beyond existing literature. We will be specifically concerned with GWs propagation (as opposed to generation) and will therefore be allowed to neglect the non-linearities crucial in the strong-gravity GWs generation regime. Our analysis is focussed on the traceless symmetric part of the two tensor sectors: it will not include the scalar and the (typically decaying) vector degrees of freedom.

We are going to address the question of how the coupling between different modes can affect the graviton propagation in a regime outside a late time de Sitter era, which is the one usually considered when studying oscillation effects in bigravity [316, 317]. This question is particularly relevant for LISA since GWs emitted from standard sirens at large redshift can probe phases of cosmological expansion that are not captured by a pure de Sitter space approximation.

Bigravity is described by the action

$$S = \int d^4x \left\{ \kappa M_{\rm Pl}^2 \sqrt{-\tilde{g}}\, \tilde{R} + \sqrt{-g} \left[M_{\rm Pl}^2 \left(R - 2m^2 V \right) + \mathcal{L}_{\rm matt} \right] \right\} , \tag{5.3.1}$$

with $g_{\mu\nu}$ and $\tilde{g}_{\mu\nu}$ the two metric tensors, $M_{\rm Pl}^2$ and $\kappa M_{\rm Pl}^2$ the corresponding squares of Planck masses and m the graviton mass. Matter is coupled only to the first metric. The interaction potential between the two metrics is indicated by V, and it takes the form

$$V = \sum_{n=0}^{4} a_n V_n , \tag{5.3.2}$$

with a_n dimensionless parameters, and

$$V_0 = 1 , \tag{5.3.3}$$
$$V_1 = \eta_1 , \tag{5.3.4}$$
$$V_2 = \eta_1^2 - \eta_2 , \tag{5.3.5}$$
$$V_3 = \eta_1^3 - 3\eta_1\eta_2 + 2\eta_3 , \tag{5.3.6}$$
$$V_4 = \eta_1^4 - 6\eta_1^2\eta_2 + 8\eta_1\eta_3 3\eta_2^2 - 6\eta_4 , \tag{5.3.7}$$

where $\eta_i = \mathrm{tr}\left[Y^i\right]$ with $Y_\mu{}^\nu = [\sqrt{X}]_\mu^\nu$. Tensor fluctuations $h_{ij}^{(1,2)}$ are defined around two FRW line elements, associated to each one of the metrics involved:

$$ds^2 = a^2(\eta)\left(-d\eta^2 + d\vec{x}^2\right), \tag{5.3.8}$$

$$d\tilde{s}^2 = \omega^2(\eta)\left(-c^2(\eta)\,d\eta^2 + d\vec{x}^2\right), \tag{5.3.9}$$

where $c(\eta)$ is the speed of the second tensor fluctuation. In what follows the ratio of scale factors is denoted by

$$\xi(\eta) = \frac{\omega(\eta)}{a(\eta)}, \tag{5.3.10}$$

while $\mathcal{H} = a'/a$ is the Hubble parameter corresponding to the first metric. The Friedmann equation for the first metric reads

$$\frac{\mathcal{H}^2}{a^2} = \frac{8\pi G}{3}\rho + m^2\left(2a_3\xi^3 + 2a_2\xi^2 + a_1\xi + \frac{a_0}{3}\right). \tag{5.3.11}$$

The theory admit two branches of solutions, but only one describes physically interesting cosmological configurations [397]. In this branch the Bianchi identities are realized in the form

$$c(\eta) - 1 = \frac{1}{\mathcal{H}(\eta)}\frac{\xi'(\eta)}{\xi(\eta)}, \tag{5.3.12}$$

and together with Friedmann equations lead to an algebraic equation for ξ:

$$\frac{8a_4}{\kappa}\xi^2 + \frac{6a_3}{\kappa}\xi + \frac{2a_2}{\kappa} + \frac{a_1}{3\kappa}\frac{1}{\xi} = \frac{\mathcal{H}^2}{m^2a^2}, \tag{5.3.13}$$

where the coefficients a_i are the parameters of the bigravity potential (5.3.2). This information about homogeneous configurations is sufficient for the scope of this analysis. Then, the evolution of the tensor perturbations in bigravity is described by a coupled system of linearized equations for the two tensor modes h_1 and h_2 [397]

$$\left[\frac{d^2}{d\eta^2} + \begin{pmatrix} 2\mathcal{H} & 0 \\ 0 & 2\left(\mathcal{H} + \frac{\xi'}{\xi}\right) - \frac{c'}{c} \end{pmatrix}\frac{d}{d\eta} + \begin{pmatrix} 1 & 0 \\ 0 & c^2 \end{pmatrix}k^2 + m^2a^2 f_1\begin{pmatrix} 1 & -1 \\ -\frac{c}{\kappa\xi^2} & \frac{c}{\kappa\xi^2} \end{pmatrix}\right]\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = 0, \tag{5.3.14}$$

where we have dropped the tensorial indices since the propagation is the same for each transverse-traceless polarization. The constant κ controls the relative size of the strength of gravitational interactions in the two sectors, while m sets the scale of the bare graviton mass. The quantity $f_1(\eta)$ is a cubic function in $\xi(\eta)$ that depends on the bigravity parameters a_i:

$$f_1(\eta) = 2\,\xi^2(\eta)\,\left[3\,a_3\,c(\eta)\,\xi(\eta) + a_2\,(c(\eta)+1)\right] + a_1\,\xi(\eta). \tag{5.3.15}$$

The time-dependent coefficients in (5.3.14) are controlled by the background solutions. We will consider the branch of solutions where scalar and vector modes are not strongly

coupled. In this branch, the propagation speed of the second tensor h_2 is determined by (5.3.12) and in what follows we assume that matter only couples with the tensor perturbation h_1. Equation (5.3.14) implies that the cosmological propagation of GWs in bigravity is characterized by three distinctive effects:

(i) the two tensor perturbations $h_{1,2}$ propagate at different speeds ($c \neq 1$ if $\xi' \neq 0$),

(ii) they have different friction terms,

(iii) they mix due to the non-diagonal mass matrix.

In general, given the time dependence of the parameters in (5.3.14), the propagation of GWs can not be solved analytically, and the system of two tensor modes cannot be diagonalized. On the other hand, in our case we can exploit the fact that the frequency of the GW ($f \sim 10^{-2}$Hz in the LISA band) is much larger than the universe expansion rate, $\mathcal{H}_0 \sim 10^{-18}$Hz. Thus, the time variation of the parameters is small compared to the frequency of the GW and we can make use of a WKB expansion to obtain approximate analytical solutions for the tensor dynamics as studied in section 5.1.1. That said, we nevertheless emphasize that for the range of redshifts probed with LISA standard sirens, $z \sim 1-5$, there is *always* some time dependence left in the parameters introduced by the scale factor $a(z)$. This is in contradistinction to LIGO sources at $z \ll 1$, for which the approximation $a \simeq 1$ can be consistently adopted, as done in previous analysis [316,317].

The WKB solution for the system of equations (5.3.14) can be obtained as follows. First, it is convenient to absorb the cosmic friction term and define a vector containing the two tensor modes,

$$\Phi = \begin{pmatrix} a h_1 \\ a h_2 \end{pmatrix} . \tag{5.3.16}$$

Then, we can express the evolution equation (5.3.14) for h_1 and h_2 in bigravity in matrix notation as

$$\left[\frac{d^2}{d\eta^2} + \hat{\nu} \frac{d}{d\eta} + \hat{C} k^2 + \hat{M} - (\mathcal{H}^2 + \mathcal{H}') \hat{I} - \mathcal{H} \hat{\nu} \right] \Phi = 0 \,, \tag{5.3.17}$$

where we have defined

$$\hat{\nu} = \begin{pmatrix} 0 & 0 \\ 0 & 2\xi'/\xi - c'/c \end{pmatrix}, \quad \hat{C} = \begin{pmatrix} 1 & 0 \\ 0 & c^2 \end{pmatrix} , \quad \hat{M} = m^2 a^2 f_1 \begin{pmatrix} 1 & -1 \\ -c/(\kappa \xi^2) & c/(\kappa \xi^2) \end{pmatrix} \tag{5.3.18}$$

using the same notation that in section 5.1.1. Next, we introduce a dimensionless expansion parameter ϵ, rescaling time as $d\eta \to d\eta/\epsilon$, and controlling different orders in a high frequency WKB approximation. We make the WKB ansatz for the solution Φ for tensor modes, i.e. Eq. (5.1.5), and aim to solve order by order in the expansion parameter ϵ. In Eq. (5.1.5) we expand over a basis of eigenvectors controlled by the matrix $\hat{E}$, associated with the matrix of eigenfrequencies $\hat{\theta}$, which appears at the exponent of (5.1.5) and is

diagonal by definition. At leading order ϵ^0 of our expansion parameter, we obtain the eigenfrequencies

$$\theta^2_{1,2} = \frac{1}{2}\Big((1+c^2)k^2 + m^2a^2f_1\left(1+\frac{c}{\kappa\xi^2}\right) - 2\mathcal{H}^2 \\ \mp\sqrt{4m^4a^4f_1^2\frac{c}{\kappa\xi^2} + \left(\left(1-c^2\right)k^2 + m^2a^2f_1\left(1-\frac{c}{\kappa\xi^2}\right)\right)^2}\Big)\,, \tag{5.3.19}$$

and the matrix of eigenvectors

$$\hat{E} = \begin{pmatrix} 1 & \frac{m^2a^2f_1}{k^2+m^2a^2f_1-\mathcal{H}^2-\theta_2^2} \\ \frac{m^2a^2f_1c/(\kappa\xi^2)}{c^2k^2+m^2a^2f_1c/(\kappa\xi^2)-\mathcal{H}^2-\theta_1^2} & 1 \end{pmatrix}. \tag{5.3.20}$$

At next to leading order, ϵ^1, the amplitude Φ_0 can be obtained solving

$$2\hat{E}\hat{\theta}\hat{G}\Phi_0' + \left(\hat{E}\hat{\theta}' + 2\hat{E}'\hat{\theta} + i\mathcal{H}'\hat{E} + i\mathcal{H}\hat{\nu}\hat{E} + \hat{\nu}\hat{E}\hat{\theta}\right)\hat{G}\Phi_0 = 0\,, \tag{5.3.21}$$

where we defined for convenience the matrix $\hat{G} \equiv e^{i\int\hat{\theta}d\eta}$. For general time dependent coefficients, this matrix equation cannot be solved analytically (because the matrices in the parenthesis should commute at any given time). However, within the regime of the WKB, a matrix exponential solution is a very good approximation (that we have checked numerically)

$$\Phi_0 = \hat{\theta}^{-1/2}\exp\left[-\frac{1}{2}\int d\eta\hat{G}^{-1}\hat{\theta}^{-1/2}\hat{E}^{-1}\left(2\hat{E}' + i\mathcal{H}'\hat{E}\hat{\theta}^{-1} + i\mathcal{H}\hat{\nu}\hat{E}\hat{\theta}^{-1} + \hat{\nu}\hat{E}\right)\hat{\theta}^{1/2}\hat{G}\right]\bar{C}_0\,, \tag{5.3.22}$$

where $\vec{C}_0$ is a vector of constant coefficients to be fixed with the initial conditions. Recall that $\hat{\theta}$ is a diagonal matrix and thus the term $\hat{\theta}^{-1/2}$ in (5.3.22) is the usual WKB scaling. If there is time dependence, there can be corrections to this scaling, which corresponds to the matrix exponential. The fact that the matrix exponential works as a solution is because the matrix in the exponent is small in this regime and, as a consequence, corrections arising from commutators of this matrix are further suppressed. For higher order corrections in the WKB expansion, one can proceed iteratively and solve Φ_1 at order ϵ^2 using the solution of Φ_0.

Having an approximate analytical solution allows us to understand the role of each parameter. In particular, we note that the speed of the massive mode has a key role in the mixing. This is better seen in the high-k limit where the phases tend to

$$\theta_1^2 = k^2 + m^2a^2f_1 - \mathcal{H}^2 + \mathcal{O}(k^{-2})\,, \tag{5.3.23}$$

$$\theta_2^2 = c^2k^2 + \frac{m^2a^2f_1c}{\kappa\xi^2} - \mathcal{H}^2 + \mathcal{O}(k^{-2})\,. \tag{5.3.24}$$

Focusing on the non-diagonal terms of the matrix of eigenvectors,

$$\hat{E}_{12} = \frac{m^2 a^2 f_1}{(c^2-1)k^2 + (1 - c/(\kappa\xi^2))m^2 a^2 f_1 + \mathcal{O}(k^{-2})}\,, \tag{5.3.25}$$

$$\hat{E}_{21} = \frac{m^2 a^2 f_1 c/(\kappa\xi^2)}{(c^2-1)k^2 + (1 - c/(\kappa\xi^2))m^2 a^2 f_1 + \mathcal{O}(k^{-2})}\,, \tag{5.3.26}$$

we notice that in the high-k limit the mode mixing is suppressed, i.e. $\hat{E}_{12} \ll 1$ and $\hat{E}_{21} \ll 1$, if each tensor propagate at a different speed, $c \neq 1$, and there is a large hierarchy between the mass term m and the wavevector k.

Altogether, the WKB solution could be summarized by the following expression

$$a(\eta)\, h_1(\eta) = \left[c_1 \bar{\phi}_1(\eta) + c_2 \hat{E}_{12}(\eta)\bar{\phi}_2(\eta) e^{i\int \delta\theta(\eta)d\eta}\right] e^{i\int \theta_1(\eta)d\eta}\,, \tag{5.3.27}$$

$$a(\eta)\, h_2(\eta) = \left[c_1 \hat{E}_{21}(\eta)\bar{\phi}_1(\eta) e^{-i\int \delta\theta(\eta)d\eta} + c_2 \bar{\phi}_2(\eta)\right] e^{i\int \theta_2(\eta)d\eta}\,. \tag{5.3.28}$$

Here, $\bar{\phi}_{1,2}$ denote components of a vector controlling the mode amplitudes. The quantities $c_{1,2}$ are constant fixed by initial conditions while $\delta\theta \equiv \theta_2 - \theta_1$. The possible mixing between h_1 and h_2 is controlled by the non-diagonal elements of the matrix of eigenvectors, i.e. $\hat{E}_{12}$ and $\hat{E}_{21}$. These entries are non-vanishing whenever the mass matrix in (5.3.14) is non-diagonal. On the other hand, we find that the size of the mixing is controlled by the relative difference of the velocities of each mode. In particular, for the mixing not to be suppressed in the regime of large k, one needs to require

$$\left(c^2 - 1\right) k^2 \lesssim m^2 a^2\,, \tag{5.3.29}$$

which follows directly from the analytic expression of $\hat{E}$ in Eq. (5.3.25). Therefore, this inequality determines the detectability of the small and large mass limits.

In the *small mass regime*, with a mass parameter of the order of the Hubble constant $m \sim \mathcal{H}_0$, the energy density proportional to m^2 in the right hand side of the Friedmann equation (5.3.11) is of the same order of magnitude of the observed vacuum energy. Then, the viable branch of solutions satisfies $\xi \ll 1$, which gives the following value for the speed of the second tensor modes [398]

$$c - 1 = 3(w+1) + \mathcal{O}\left(\frac{m^2}{G\rho}\right)^2\,. \tag{5.3.30}$$

This implies that during matter domination $c^2 \sim 16$, and during radiation domination $c^2 \sim 25$. Therefore, given that LISA frequency of GWs is much larger than the rate of expansion, $k_{\rm LISA} \sim 10^{16}\mathcal{H}_0$, inequality (5.3.29) is not satisfied and mixing among different modes is negligible in the small mass regime. Moreover, this mass range is also still far from being constrained through the modified dispersion relation with LISA. In particular, the bound on an effective mass term $m\, f_1^{1/2}$ in this case is [355]

$$m\, f_1^{1/2} \lesssim 10^{-26}\,\frac{\rm eV}{c^2}\left(\frac{10\,{\rm Gpc}}{D}\,\frac{f}{10^{-2}\,{\rm Hz}}\,\frac{100}{\rm SNR}\right)^{1/2}\,, \tag{5.3.31}$$

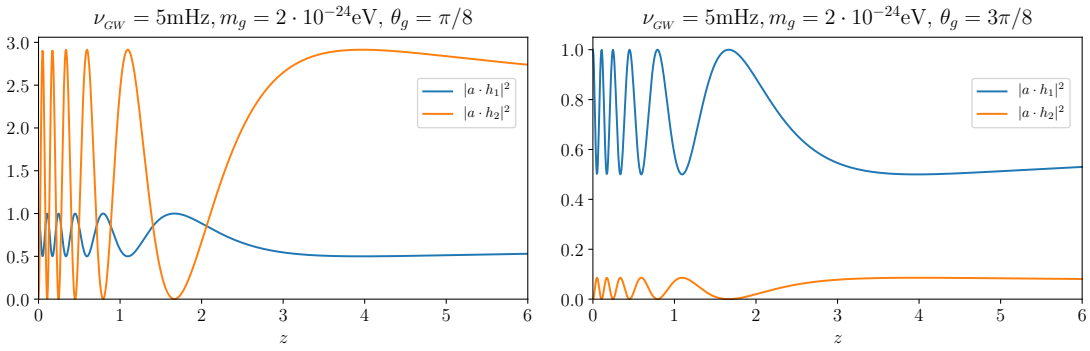

Figure 5.10. Amplitude of the tensor perturbations h_1 and h_2 for a GW emitted with frequency $\nu_{GW} = 5\text{mHz}$. The left and right panels correspond to different mixing angles for the same mass $m_g = 2 \cdot 10^{-24}\text{eV}$ and we have normalised the initial amplitude of h_1 to 1. See eq (5.3.34) for the definitions of m_g and θ_g.

where we have introduced the expected distance, frequency and signal-to-noise ratio (SNR $\sim 1/(f\Delta t)$) of a massive BH binary in the LISA band. This is seven orders of magnitude larger than $\mathcal{H}_0$, thus far from the small mass regime.

On the contrary, in the *large mass regime*, when the inequality (5.3.29) is satisfied, one can have large mixings among modes with interesting phenomenological consequences, such as graviton oscillations. In this regime ξ is approximately constant and, as a consequence, $c \simeq 1$ (recall (5.3.12)). In particular, we find

$$(c^2 - 1) \lesssim (\mathcal{H}^2 - \mathcal{H}_0^2\Omega_\Lambda)/(m^2a^2)\,, \tag{5.3.32}$$

where Ω_Λ is the density of DE. As a consistency check, we see that in the pure de Sitter limit, the speed is exactly luminal, $c_{dS} = 1$. The condition to have mode mixing, eq (5.3.29), is satisfied for

$$m^4a^4 \gtrsim k^2(\mathcal{H}^2 - \mathcal{H}_0^2\Omega_\Lambda)\,. \tag{5.3.33}$$

For LISA frequencies this leads to a bound $m \gtrsim 10^8\mathcal{H}_0$, which improves by a couple of orders of magnitude the LIGO detectability range $m \gtrsim 10^{11}\mathcal{H}_0$ [317] (recall that $\mathcal{H}_0 \sim 10^{-33}\text{eV}$). We stress here that in the large mass regime the evolution of the universe cannot reproduce the observed accelerated expansion unless we include an effective cosmological constant term compensating the large mass in the Friedmann equation. This means that, by restricting the analysis to the large mass regime, the present work does not fully probe bigravity: it does not capture the region of parameter space where $m \sim \mathcal{H}_0$, which is the most propitious for self-acceleration. On the other hand, our analysis can also serve as a proxy for other scenarios supporting oscillations in the luminosity distance at late time. A case in point are models with vector gauge fields [319, 396].

We henceforth restrict the analysis to the regime of (5.3.33) and investigate its phenomenological consequences. It is convenient to define an effective mass m_g and a mixing

angle θ_g as in section 5.1.2,

$$m_g^2 \equiv m^2 f_1\left(\frac{1}{\kappa\,\xi^2}+1\right) \quad \text{and} \quad \theta_g \equiv \tan^{-1}\left[\sqrt{\kappa\,\xi^2}\right]. \tag{5.3.34}$$

Whenever $(c^2-1)\,k^2 < m_g^2 a^2$, one can expand the phases associated to each tensor mode as

$$\theta_1^2 = k^2\left(1+\frac{(c^2-1)\,\kappa\,\xi^2}{1+\kappa\xi^2}\right) - \mathcal{H}^2 + \mathcal{O}\left(\frac{(c^2-1)^2\,k^4}{m_g^4 a^4}\right), \tag{5.3.35}$$

$$\theta_2^2 = k^2\left(1+\frac{(c^2-1)}{1+\kappa\xi^2}\right) + m_g^2 a^2\left(1+\frac{c^2-1}{2(1+\kappa\xi^2)}\right) - \mathcal{H}^2 + \mathcal{O}\left(\frac{(c^2-1)^2\,k^4}{m_g^4 a^4}\right). \tag{5.3.36}$$

Analogously, the eigenvectors simplify to

$$\hat{E} = \begin{pmatrix} 1 & -\kappa\xi^2\left(1-\left(\frac{(c^2-1)k^2}{m_g^2a^2}+\frac{(c^2-1)}{2}\right)\right) \\ 1-\frac{(c^2-1)k^2}{m_g^2a^2} & 1 \end{pmatrix} + \mathcal{O}\left(\frac{(c^2-1)^2\,k^4}{m_g^4a^4}\right). \tag{5.3.37}$$

In order to estimate how large is the correction to the GW amplitude w.r.t. the $c=1$ case, we can parametrize the effective mass with a dimensionless constant β via $(m_g\,a) \sim \beta\cdot(k\,\mathcal{H}_0)^{1/2}$, which controls our complying with the large mass regime defined in (5.3.33). It follows that the largest correction to the matrix of eigenvectors (5.3.37) scales with β^{-4}. Therefore, provided that $\beta \gtrsim 5$, we can neglect this correction in the amplitude.

The mixing among modes affect the tensor speed and the luminosity distance. The modification in the propagation speed of the lightest tensor h_1, defined as $\alpha_T \equiv c_T^2 - 1$, is given by

$$\alpha_T = \frac{(c^2-1)\;\kappa\xi^2}{(1+\kappa\xi^2)}, \tag{5.3.38}$$

and scales as $\alpha_T \sim \beta^{-2}(k/\mathcal{H}_0)^{-2}$. For LISA, α_T is smaller than 10^{-16} in the large mass limit. However, it might still be observable given that this is a cumulative effect over long travel distances. A prompt EM counterpart can give constraints of the order of

$$\alpha_T \lesssim 2\times 10^{-17}\left(\frac{10\text{Gpc}}{D}\right)\left(\frac{\Delta t}{10\text{s}}\right), \tag{5.3.39}$$

where Δt is the difference in the time of arrival and D the distance to the source.

Next, we study how the mixing of the tensor modes $h_{1,2}$ can leave an imprint in the GW luminosity distance. Focussing on the regime $(m_g\,a) \gtrsim 5\cdot(k\,\mathcal{H}_0)^{1/2}$, the only time dependence in the amplitude is introduced by the scale factor $a(\eta)$. Using the definitions (5.3.34) of the effective mass m_g and the mixing angle θ_g, the amplitude of the tensor component h_1 is

$$|a\,h_1|^2 = h_0^2\cos^4\theta_g\left(1+\frac{\tan^4\theta_g}{\theta_2/\theta_2(\eta_e)}+\frac{2\tan^2\theta_g}{\sqrt{\theta_2/\theta_2(\eta_e)}}\cos\left[\int_{\eta_e}^{\eta}\delta\theta d\eta'\right]\right), \tag{5.3.40}$$

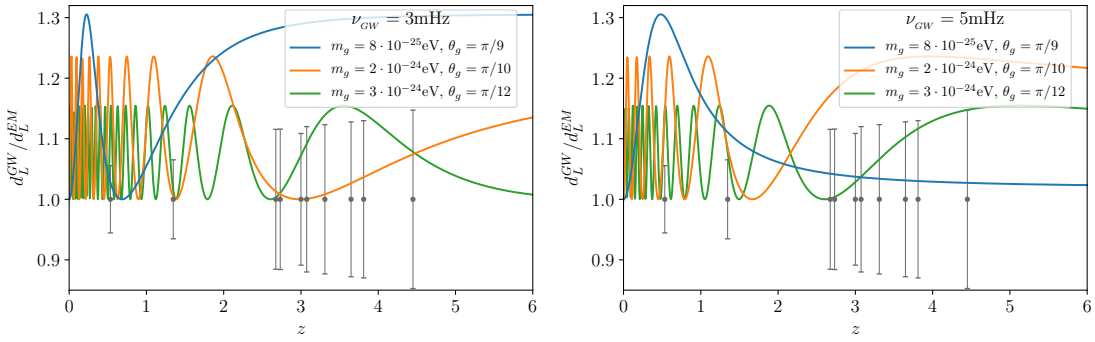

Figure 5.11. GW luminosity distance in the large mass limit of bigravity for different values of m_g and θ_g. To visualize the range of masses and mixing angles that could be probed with LISA, we have included the errors of a representative catalog of massive black-hole binaries which could be used as standard sirens (see details in [8]).

where h_0 is the amplitude in the $\theta_g = 0$ case, quantities depend on the time η, and η_e is the time of emission. Also, $\theta_1^2 = k^2 - \mathcal{H}^2$, $\theta_2^2 = k^2 + m_g^2 a^2 - \mathcal{H}^2$, $\delta\theta = m_g^2 a^2/(2k) + \mathcal{O}(k^{-3})$. As a result, if $m_g^2 a^2 \sim k\mathcal{H}_0$, the oscillation frequency is of order $\mathcal{H}_0$. A similar expression can be obtained for $|a\, h_2|$. One should note that while $|a\, h_1|$ is never larger than the initial value h_0, for $\theta_g < \pi/4$, h_2^2 can exceed h_0^2 (up to $4h_0^2$). In the opposite limit, $\theta_g > \pi/4$, h_2^2 is always less than h_0^2. This could be observed by comparing the left and right panels of Fig. 5.10, where the mass is fixed but the mixing angles vary.

We can now compare the luminosity distance of GWs in bigravity with the one of EM radiation, $d_L^{\rm gw}/d_L^{\rm em}$. We focus on the amplitude of the lightest tensor mode, h_1, which is the one we assume to be coupled to matter. From (5.3.40), we obtain

$$\frac{d_L^{\rm gw}}{d_L^{\rm em}} \simeq \frac{1}{\cos^2\theta_g}\left(1 + \frac{\tan^4\theta_g}{\theta_2/\theta_2(\eta_e)} + \frac{2\tan^2\theta_g}{\sqrt{\theta_2/\theta_2(\eta_e)}}\cos\left[\int_{\eta_e}^{\eta}\frac{m_g^2 a^2}{2k}d\eta'\right]\right)^{-1/2}, \qquad (5.3.41)$$

which is equivalent to equation (5.2.3). We see that this ratio can become larger than one and display oscillatory patterns. Possible configurations range in principle[8] between $\theta_g = 0$ (corresponding to $\kappa\xi^2 = 0$) and $\theta_g = \pi/2$ (corresponding to $\kappa\xi^2 \to \infty$), and the maximum mixing occurs at $\theta_g = \pi/4$ ($\kappa\xi^2 = 1$). Moreover, in the limit in which $\theta_2/\theta_2(t_e) = 1$ (which is a good approximation in the high k limit), $d_L^{\rm gw}/d_L^{\rm em}$ is symmetric around $\pi/4$, i.e. $h_1^2(\pi/4 - \varphi) = h_1^2(\pi/4 + \varphi)$.

Given that the ratio (5.3.41) oscillates as a function of redshift, a two-parameter phenomenological parametrization such as (4.2.3) is not expected to perform well in this case, and an analysis specific for this model is necessary. We plot the oscillatory behavior of the GW luminosity distance for different masses and mixing angles in Fig. 5.11. For

[8]It is actually not possible to push the theory to the asymptotic regions for the following reasons: (i) $\kappa \to 0$ corresponds to a vanishing kinetic term for the second metric so that it becomes infinitely strongly coupled; (ii) taking the limit $\kappa \to \infty$, the second metric decouples and gets effectively frozen to a fixed background value.

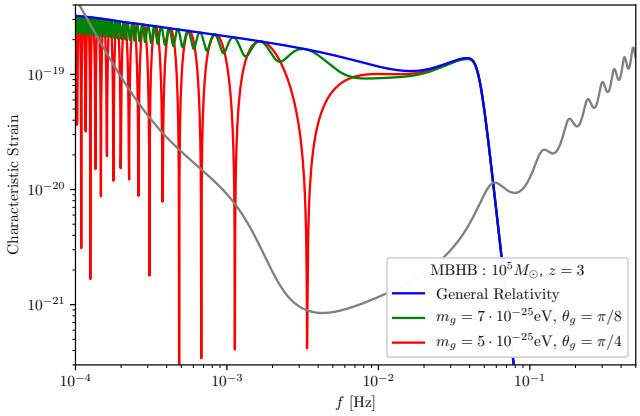

Figure 5.12. GW strain versus frequency in bigravity for a signal from a massive BH binary (MBHB) merger event detectable with LISA (in grey the LISA sensitivity curve). We compare the signal strain for two bigravity sets of parameters (fixed mass, changing mixing angle) with respect to GR. Notice the characteristic oscillations in frequency of the GW strain for this theory.

a qualitative analysis, we include error bars in the measurements of luminosity distances for a representative LISA catalogue of massive black-hole binaries acting as standard sirens at high-redshift. For details on the astrophysical modeling of this catalog, we refer to the details presented in [8].

Assuming a fixed cosmology, we can compare the high-mass prediction for the luminosity distance ratio (5.3.41) with LISA catalogs and perform a χ^2 analysis. Oscillation effects can be observed for masses $m_g \gtrsim 2 \cdot 10^{-25}$ eV. The amplitude of the oscillation, and thus its detectability, increases as θ_g approaches to $\pi/4$ from above or below. The mixing angle range where oscillations can be detected is $0.05\pi \lesssim \theta_g \lesssim 0.45\pi$. LISA will provide a ~ 3 order of magnitude improvement in mass sensitivity over the current LIGO limit, which probes $m_g \gtrsim 10^{-22}$eV [317], due to the larger oscillation baseline and the lower detection frequency. We conclude that the use of standard sirens will strengthen existing bounds towards smaller values for the mass.

However, one should note that these bounds are based on standard sirens with a single fiducial frequency, i.e. a monochromatic GW. Since GW oscillation effects depend strongly on the frequency, a coallescing binary would experience a time-dependent modulation of the amplitude as the orbital frequency increases, leading to a distinct signal. Remarkably, no electromagnetic counterparts are necessary to study such effects. As a representative example, we plot in Fig. 5.12 the strain of a GW signal from a massive BH binary as a function of frequency and its modification in bigravity. The GW oscillations lead to a distinctive frequency profile, which would be interesting to further characterise in order to establish to what extent it can be probed with LISA. Moreover, the initial wave packet emitted might decohere while traveling. In that case the event would be followed by an "echo" signal, which may be detectable if the mixing is sufficiently large [318]. The inclusion of decoherence effects and frequency dependence would provide further means to test GW oscillations. We leave this to future work.

Part III

Black-holes in the Early Universe

CHAPTER

6

Primordial black-hole formation

The tale of modern cosmology is rather elegant. Large scale structures of the universe are seeded by quantum fluctuations generated during the first instants and stretched to cosmological scales by the exponential expansion of inflation. Yet, quantum fluctuations could also have a decisive role in the small scales of the universe. This would be the case if there was an enhancement of the power spectrum at small scales, triggering the collapse of large fluctuations into primordial black-holes (PBH) [399]. PBHs are interesting objects because they leave imprints throughout the history of the universe by their energy injections, dynamical effects and GWs, topics reviewed recently in [28, 226, 400, 401]. Depending of their masses, they might also be the BHs detected by LIGO [239–241] or the seeds for supermassive BHs [224, 225]. Moreover, as we have introduced in Sec. 1.2, if they are abundant enough, they could constitute a fraction of the dark matter. But, probably above all, PBHs could open a unique window into the very early Universe.

PBHs can be generated by many different mechanisms. We will focus on PBHs formed from large, primordial curvature perturbations. In that case, the fraction of PBHs formed in the early universe β_f is determined by the probability that a given primordial curvature fluctuation ζ is above a certain threshold

$$\beta_f(M) = \int_{\zeta_c}^{\infty} P(M;\zeta) d\zeta. \tag{6.0.1}$$

Therefore, the abundance of PBHs is both sensitive to the probability density function (PDF) $P(M;\zeta)$ and the value of the threshold ζ_c. Here, we have expressed the PDF in terms of the mass M of the PBHs formed instead of the number of e-folds. This is because the mass of the PBHs is related to the size of the causal horizon collapsing that, at the same time, depends on the time of formation (see e.g. [399])

$$M_{\rm PBH} \simeq 4\pi\gamma \frac{M_{\rm pl}^2}{H_{\rm inf}} e^{2N}\,, \tag{6.0.2}$$

where $M_{\rm pl}$ is the reduced Planck mass, $H_{\rm inf}$ the energy scale of inflation and γ efficiency parameter encapsulating the details of the gravitational collapse and the efficiency of reheating, that we fix to be $\gamma \sim 0.2$.

With the fraction of PBHs forming $\beta_f(M)$ one can compute the contribution of PBHs to the energy density of the universe and determine the fraction of the DM that

they represent today. In this sense, it is convenient to compute this at the time of matter-radiation equality

$$\Omega_{\rm PBH}^{\rm eq} = \int_{M_{\rm ev}}^{M_{\rm eq}} \beta_{\rm eq}(M) d\ln M \tag{6.0.3}$$

where $M_{\rm ev}$ is the mass threshold for Hawking evaporation and we have assumed that the fraction of PBH at the time of equality has grown due to cosmic expansion from the time of formation $\beta_{\rm eq}(M) = e^{(N_{\rm eq}-N_f)}\beta_f(M)$.

The PDF is governed by the physics in the early universe while the threshold is subject to the conditions at the time of formation. The analysis of the gravitational collapse of curvature perturbations to form PBHs and the appropriate threshold condition has been an active line of research in the past years [402–404], with recent analysis on the dependence of the threshold in the shape of the power spectrum [405, 406] and on the non-linear relation between the curvature perturbations and the density contrast [407, 408]. Note that the threshold is also sensitive to the equation of state (EoS) at the time of formation. For example, the QCD phase transition makes the EoS to drop, increasing the production of PBHs of $\mathcal{O}(1M_\odot)$ [409, 410]. Since we are interested in the physics of the early universe that enhances the PDF rather than the physics at the time of formation, we will follow the above prescription for the formation of PBHs with a fixed ζ_c. Using another prescription might change the numbers for β_f but would not change the results qualitatively.

Large curvature fluctuations can be produced by very different means in the early universe. We will be interested in producing them during inflation and, in particular, we will focus on single-field models. First, we will investigate how to produce an enhancement of the power spectrum $\mathcal{P}_\zeta$ in section 6.1. If the curvature perturbations were Gaussian, then the PDF would be described solely by its variance $\sigma^2 = \mathcal{P}_\zeta(k_M)$ and the fraction of PBH at formation would simplify to

$$\beta_f(M) = \int_{\zeta_c}^{\infty} \frac{{\rm d}\zeta}{\sqrt{2\pi}\sigma} e^{-\frac{\zeta^2}{2\sigma^2}} = \frac{1}{2}{\rm erfc}\left(\frac{\zeta_c}{\sqrt{2}\,\sigma}\right), \tag{6.0.4}$$

where k_M is the wave-number associated to the scale in which PBH of mass M form. Thus, a larger power spectrum would increase the fraction of PBHs. In terms of the density contrast, for a Gaussian distribution, the relevant quantity is

$$\sigma_\delta^2 = \int_0^\infty \frac{dk}{k} \mathcal{P}_\delta(k) W^2(k\,R)\,, \tag{6.0.5}$$

where $\mathcal{P}_\delta(k)$ is the density power spectrum, $W(k\,R)$ is the smoothing window function and R the horizon scale. The fraction is then $\beta(M) = {\rm erfc}\left(\delta_c/\sqrt{2}\sigma_\delta\right)/2$ and the reasoning with the power spectrum follows similarly.

Such enhancement of $\mathcal{P}_\zeta$ can occur if there is a second plateau in the potential [411]. As we will discuss in section 6.2, in this class of models, in order to have sufficiently large fluctuations, the inflationary dynamics has to deviate from slow-roll (SR) [412–414]. In the next chapter 7, we will present a particular inflationary model based on critical Higgs inflation that produces PBHs through this mechanism [4]. Other examples of single-field

models producing PBHs are double inflation [412], radiative plateau inflation [415] or some string realizations [416–418].

In order to understand the dynamics of the inflaton in these scenarios, it is convenient to parametrize the evolution in terms of the Hubble flow and its derivatives with respect to the number of e-folds $dN = Hdt$. For that purpose, we use the following Hubble-flow parameters

$$\epsilon_n = \frac{d\ln|\epsilon_{n-1}|}{dN}\,, \tag{6.0.6}$$

where the initial parameter in the series corresponds to the inverse Hubble parameter $\epsilon_0 = 1/H$. Then, the first parameter is $\epsilon = -H(N)'/H$. Within this language, slow-roll inflation is defined by $|\epsilon_n| \ll 1$.

Importantly, this period beyond SR necessary to produce enough PBHs leads to an exponential growth of the quantum modes [419, 420]. The quantum back-reaction of these modes can alter the abundance of PBHs by modifying the power spectrum and introducing non-Gaussian corrections [6]. We will study this quantum diffusion effects in detail in chapter 8.

6.1 Enhancement of the power spectrum

In order to obtain the primordial power spectrum of inflation we need to solve the evolution of the inflaton and metric quantum fluctuations. For that purpose, we use the gauge invariant curvature fluctuation

$$\zeta = \frac{u}{z} = \frac{a\delta\phi}{z}\,, \tag{6.1.1}$$

where u is the Mukhanov-Sasaki variable [421, 422] and we define the classical field $z = a\,d\phi/dN$. Note that we are absorbing the scalar metric perturbation Φ in the inflaton fluctuations $\delta\phi$ by adding a correction to the time dependent, effective mass term. This is equivalent to work in the uniform-curvature gauge.

In Fourier space, the evolution of the curvature perturbations $\zeta_k = u_k/z$ is described by,[1]

$$\frac{d^2\zeta_k}{dN^2} + (3 - \epsilon + \epsilon_2)\,\frac{d\zeta_k}{dN} + \left(\frac{k}{aH}\right)^2 \zeta_k = 0\,, \tag{6.1.2}$$

where we have assumed that the sound speed is constant.[2] This equation has two well-defined regimes delimited by the comoving horizon size $d_{\rm H} = 1/aH$. At sub-horizon scales ($k \gg aH$), the friction term is irrelevant and the equation describes a free field in

[1] This can be easily derived from the usual Mukhanov-Sasaki equation $u_k''(\eta) + (k^2 - z''/z)u_k = 0$ by changing from conformal time to the number of e-folds, $dN = \mathcal{H}d\eta$. Also, we have used that $d\ln z/dN = 1 + \epsilon_2/2$.

[2] Note that a varying sound speed $c_s(\eta)$ can also serve to enhance the power spectrum as recently studied in [423].

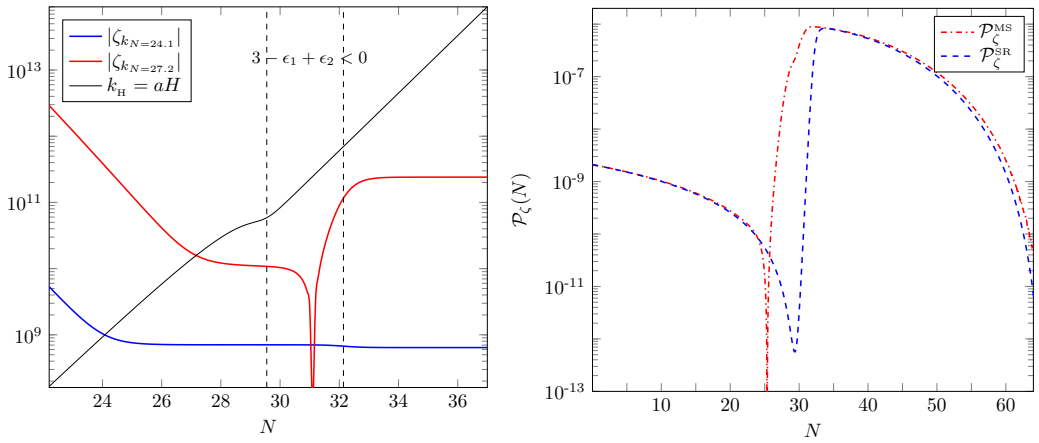

Figure 6.1. On the left, evolution of two curvature perturbations ζ_k that exit the horizon $k = aH$ at different times. There is an enhancement of the modes after horizon crossing in the region between dashed lines where $3 - \epsilon + \epsilon_2 < 0$ is satisfied. On the right, effect of this enhancement in the power spectrum, comparing the SR approximation in $\mathcal{P}_\zeta$ (6.1.6), blue dashed line, with the numerical solution of the Misao-Sasaki equation (6.1.5), red dash-dotted line.

Minkowski space, which can be normalized as a Bunch-Davis vacuum. At super-horizon scales ($k \ll aH$), the asymptotic solution is

$$\left.\frac{d\zeta_k}{dN}\right|_{k\ll aH} = C_2 e^{-\int(3-\epsilon+\epsilon_2)dN} = \tilde{C}_2 e^{-3N+\ln H-\ln\epsilon}\,, \tag{6.1.3}$$

implying that there are two modes: one constant and another evolving

$$\zeta_{k\ll aH} = C_1 + \tilde{C}_2 \int e^{-3N+\ln H-\ln\epsilon} dN\,. \tag{6.1.4}$$

Therefore, depending on the sign of $3 - \epsilon + \epsilon_2$ the second mode will be exponentially decaying or growing. The power spectrum associated to the curvature perturbations is then

$$\mathcal{P}_\zeta^{\rm MS} = \left.\frac{k^3}{2\pi^2}|\zeta_k|^2\right|_{k\ll aH}\,, \tag{6.1.5}$$

which is obtained solving numerically the mode equation (6.1.2). Interestingly, the mode equation could be rewritten as a system of two first-order equations for the adiabatic and entropy perturbations [419], showing that the only source of growth of the curvature after horizon crossing are the entropy perturbations.

Usually this evolving mode is exponentially suppressed because ϵ and ϵ_2 are small. Then the curvature perturbation becomes constant very rapidly after horizon crossing and its power spectrum can be evaluated at this time $k = aH$. The power spectrum itself can be computed solving the equation for the curvature perturbation with the proper normalization. Doing that one arrives at the standard result

$$\mathcal{P}_\zeta^{\rm SR} \simeq \frac{\kappa^2}{8\pi^2}\frac{V(\phi)}{\epsilon(3-\epsilon)}\,, \tag{6.1.6}$$

where the $\simeq$ indicates that this is not an exact result beyond slow-roll. From this formula one can easily elucidate a method to amplify the fluctuations, i.e. slow down the inflaton to reach a small value of ϵ since the power spectrum is inversely proportional to ϵ. Note that in order to have a significant peak one needs to go beyond slow-roll [412–414] and the power spectrum (6.1.6) might not be a good approximation [414, 415].

However, if we are beyond slow-roll, the friction term could change sign and the non-adiabatic mode would grow exponentially [419, 420]. This will happen whenever $\epsilon_2 < -3 + \epsilon$ and the decaying mode will become a growing mode. Thus, the curvature perturbation ζ will grow after horizon crossing. This growth on super-horizon scales is therefore another way to enhance the power spectrum.[3] This enhancement after horizon crossing was already considered as a source of PBH in [419]. More recently, this enhancement has been shown to be important for different single-field inflationary models of PBH production [415–417, 424]. In the next section we will consider a toy model with a quasi-inflection point (7.1.6) where there will also be such growth after horizon crossing. This can be seen in the left panel of Fig. 6.1, where we have presented the evolution of two different ζ_k exiting the horizon at different times. Modes that exit the horizon near the region or in the region where $3 - \epsilon + \epsilon_2 < 0$ will suffer from this rapid growth. On the contrary, modes that exit the horizon well before or after this region will soon become constant after horizon crossing.

The growing modes outside the horizon affect the computation of the power spectrum. This can be seen graphically in the right panel of Fig. 6.1, where we compare the SR approximation for the power spectrum (6.1.6) with the numerical solution of the Misao-Sasaki equation (6.1.5). The main differences occur in two regions: when $3 - \epsilon + \epsilon_2 < 0$ and at the end of inflation. In the first region, the growth of the perturbations after horizon crossing makes the peak of $\mathcal{P}_\zeta$ to move to higher scales. Depending on the size of the peak, this could imply approaching to the detectability region of spectral distortions. Moreover, in certain scenarios such as [415], if the region $3 - \epsilon + \epsilon_2 < 0$ is long enough, the height of the peak itself could increase w.r.t. the SR approximation for $\mathcal{P}_\zeta$. In the second region, at the end of inflation, there are also differences with the SR approximation. This is because ϵ_2 is of order 1 during this period and the scalar, metric fluctuations Φ are excited. For certain shapes of the potential at the end of inflation, this difference could be significantly larger

The fact that there is a strong enhancement of quantum fluctuations during the period beyond slow-roll will have important consequences for quantum diffusion. This is because then the quantum diffusion can dominate over the classical drift. In chapter 8 we will study the back-reaction of these large quantum fluctuations on the classical inflationary dynamics using the formalism of stochastic inflation beyond slow-roll. We will see that quantum diffusion can further enhance the power spectrum. Moreover, non-Gaussian correction will also enter in the game.

[3] Nevertheless, one should note that this enhancement on super-horizon scales does not occur for the tensor perturbations [419] because their friction does not flip sign, since they follow the evolution equation $\frac{d^2 v_k}{dN^2} + (1-\epsilon)\frac{dv_k}{dN} + \left[\left(\frac{k}{aH}\right)^2 - (2-\epsilon)\right] v_k = 0$ where $v_k = a\, h_k$.

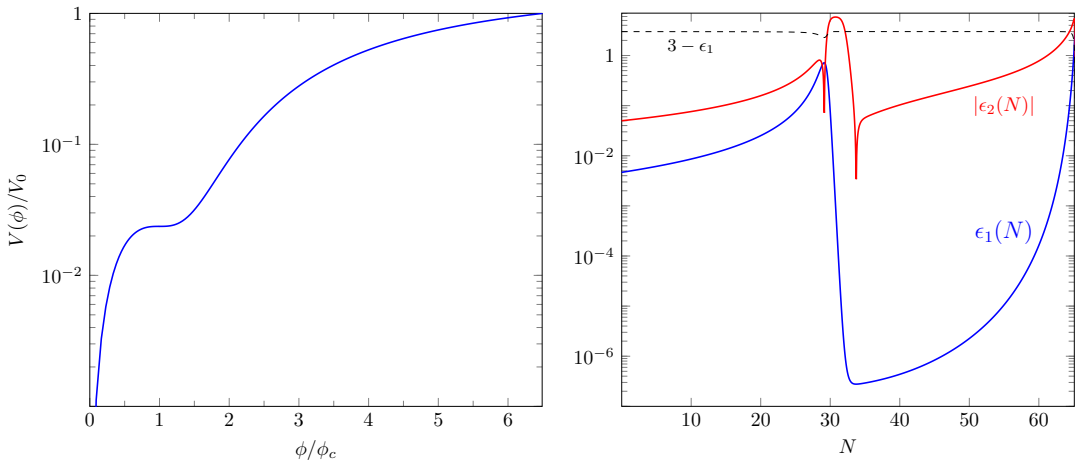

Figure 6.2. On the left, inflationary potential with a quasi-inflection point at ϕ_c for the toy model (7.1.6). On the right, evolution of the first slow-roll parameters ϵ and ϵ_2, cf. (6.0.6). The dashed line corresponds to $3-\epsilon$. In the region where $|\epsilon_2|$ is above this line, there will be an enhancement of the power spectrum after horizon crossing.

6.2 Inflation with a quasi-inflection point

We are interested in studying inflationary scenarios with an quasi-inflection point because they naturally give rise to a peak in the power spectrum [411]. The fact that there is a quasi-inflection point ϕ_c in the potential is given by the condition $V_{,\phi}(\phi_c) \approx V_{,\phi\phi}(\phi_c) \approx 0$, where $_{,\phi}$ indicates a partial derivative with respect to the inflaton ϕ. For simplicity we will describe inflation with a toy model potential [411]

$$V(\phi) = \frac{1}{12}\frac{6m^2\phi^2 - 4\alpha\phi^3 + 3\lambda\phi^4}{(1+\xi\phi^2)^2}\,, \tag{6.2.1}$$

and choose the parameters to fulfill the quasi-inflection point condition accordingly[4]. Note that we are interested in a quasi-inflection point rather than a true inflection point, $V_{,\phi} = V_{,\phi\phi} = 0$, since we want the inflaton to continue rolling ending inflation at the minimum of the potential avoiding getting trapped in the self-reproduction regime [425]. A graphical representation of this potential can be found in the left panel of Fig. 6.2. The $(1+\xi\phi^2)^2$ function in the denominator is introduced to flatten the potential at large field values, improving the agreement with CMB observations. Within the context of the ϵ_n parameters, the spectral index n_s and the tensor-to-scalar ratio r are given respectively by $n_s = 1 - 2\epsilon - \epsilon_2$, $r = 16\epsilon$. Now, before introducing the effects of quantum diffusion, let us first review the classical dynamics of this toy model and how an enhancement of the power spectrum is produced.

[4]In particular, we use in the plots $\lambda = 1$, $\xi = 2.3$, $\alpha = 6\lambda\,\phi_c/(3+\xi^2\phi_c^4) - 4.3\cdot 10^{-5}$ and $m^2 = \lambda\phi_c^2(3+\xi\phi_c^2)/(3+\xi^2\phi_c^4)$, with the field starting at $\phi_{65} = 6.5\phi_c$.

The classical dynamics of the inflaton fields is determined by its evolution equation

$$\frac{d^2\phi}{dN^2} + (3-\epsilon)\frac{d\phi}{dN} + \frac{(3-\epsilon)}{\kappa^2}\left(\ln V\right)_{,\phi} = 0\,. \tag{6.2.2}$$

At large field values the inflaton slow-rolls the potential giving rise to the CMB fluctuations. Before crossing the quasi-inflection point, the inflaton acquires some inertia that is rapidly lost when the inflaton gets closer to the quasi-inflection point and starts feeling the flatness of the potential around that point. This makes the inflaton to spend many e-folds of its evolution around this quasi-inflection point. During this time, also known as ultra slow-roll [426–429], fluctuations would be largely amplified and could lead to the production of PBH. However, for suitable potential, it retains enough inertia to cross this point and end inflation oscillating around the minimum of the potential.

This characteristic evolution has two implications. Firstly, since the inflaton slows down around the quasi-inflection point, the first Hubble-flow parameter ϵ will become very small. This will introduce an enhancement of the power spectrum since $\mathcal{P}_\zeta \propto 1/\epsilon$. Secondly, the fact that the inflaton changes its velocity very rapidly as it gets to the quasi-inflection point will bring its evolution outside of slow-roll. This can be seen by the fact that $|\epsilon_2|$ becomes large. In turn, this can produce a further enhancement of the power spectrum by exciting growing modes outside the horizon. A characterization of this specific behavior of $\epsilon(N)$ and $\epsilon_2(N)$ in these scenarios is plotted in the right panel of Fig. 6.2. 1805.03946.

CHAPTER

7

PBH production in Critical Higgs Inflation

The nature of the inflaton field responsible for the initial acceleration of the universe is still unknown. Observations of the temperature and polarization anisotropies in the cosmic microwave background (CMB) suggests a special inflaton dynamics, dominated by a flat plateau on large scales [430]. Such type of potentials arise naturally in models of Higgs Inflation [431], where the scalar field responsible for inflation is the Higgs boson of the Standard Model (SM) of Particle Physics, with its usual couplings to ordinary matter (gauge fields, quarks and leptons), plus a non-minimal coupling ξ to gravity.[1] This economical scenario passes all solar system and CMB observational constraints, predicts a small tensor-to-scalar ratio determined by the plateau of the potential at large field values leading to $r \sim 12/N^2$. Since the amplitude of the power spectrum at CMB scales ($\mathcal{P}_\zeta(k_{\rm CMB}) \sim 2 \cdot 10^{-9}$) is fixed by the ratio

$$V(\phi_{\rm CMB}) \sim \lambda/\xi^2\,, \tag{7.0.1}$$

and the SM Higgs self-coupling is $\lambda \sim 0.13$, the non-minimal coupling in this scenario must be rather large, $\xi \sim 10^5$.

In order to connect particle physics phenomena at low-energies with inflationary dynamics at high-energies, we will assume that there is no new physics between the energy scale of current particle accelerators (TeV) and the Planck scale (10^{15}TeV). In this way, both scales could be linked through the renormalization group equations (RGE), which determine how the SM couplings scale with energy. In particular, for Higgs inflation we will be interested in the running of the Higgs quartic self-coupling λ and of the non-minimal coupling to gravity ξ. The running of the SM couplings is most sensitive to the value of the strong coupling α_s, the quark top mass $m_{\rm top}$ and the Higgs mass $m_{\rm Higgs}$. This can be seen in Fig. 7.1, where we plot the running of λ and its beta function $\beta_\lambda = d\lambda/d\ln\mu$ using the 2-loop RGE, which can be found for instance in the appendix of Ref. [432]. We present λ and β_λ for different values of $m_{\rm Higgs}$ and $m_{\rm top}$ within the

[1] In the context of the scalar-tensor theories previously studied, the gravity sector of Higgs inflation belongs to a quartic Horndeski model with $G_4(\phi) = \left(\frac{1}{2\kappa^2} + \frac{\xi}{2}\phi^2\right)$ and $G_2(\phi, X) = -\frac{1}{2}(\partial\phi)^2 - \frac{1}{4}\lambda\phi^4$, whose Lagrangian was given in equation (A.1.25).

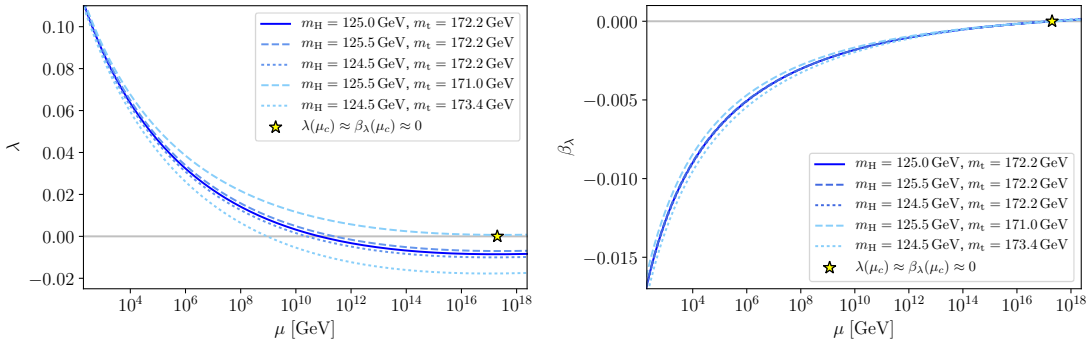

Figure 7.1. Running of the Higgs self-coupling $\lambda(\mu)$ and its beta function $\beta_\lambda(\mu)$ following the renormalization group equations of the SM at 2-loops. The 13TeV LHC central values for $m_{\rm Higgs}$ and $m_{\rm top}$ correspond to the solid line leading to a negative λ above 10^{12}TeV. Within the present uncertainties, see (7.0.2), there could be a critical point in which $\lambda(\mu_c) \approx \beta_\lambda(\mu_c) \approx 0$.

current uncertainties at 13TeV in the LHC[2]

$$m_{\rm Higgs} = 124.98 \pm 0.28\,\mathrm{GeV}\ [434]\,, \quad m_{\rm top} = 172.25 \pm 0.70\,\mathrm{GeV}\ [435]\,. \tag{7.0.2}$$

As it can be observed in Fig. 7.1, the central value of $m_{\rm Higgs}$ and $m_{\rm top}$ (solid line) leads λ to become negative at high-energies, a signature of the metastability of the electroweak vacuum [436]. Interestingly, within present uncertainties, λ could also remain positive up to near the Planck scale where its value vanishes at the same time that its beta function. In other words, there could be a *critical point* at μ_c with $\lambda(\mu_c) \approx \beta_\lambda(\mu_c) \approx 0$, where $\lambda(\mu)$ has a minimum.

If Higgs inflation takes place at the critical point, the right amplitude of CMB anisotropies can be achieved with a relatively smaller ξ coupling [437, 438]. This is because the amplitude of the fluctuations is fixed by

$$V(\phi_{\rm CMB}) \sim \lambda(\mu)/\xi^2\,. \tag{7.0.3}$$

Since λ is very small at the critical point, $\lambda(\mu_c) \sim 10^{-6}$, the non-minimal coupling may be $\xi \sim 10 - 100$. As a bonus, the tensor-to-scalar ratio can be pushed to larger values since most of the e-folds are spent near the critical point. Note that this scenario only considers the running of λ and takes ξ to be constant, neglecting its running. This assumption follows from the 1-loop running of the SM non-minimally coupled to gravity, whose beta function reads [439]

$$\beta_\xi = \frac{\partial \xi}{\partial \log \mu} = -\frac{1}{16\pi^2}\left(\xi - \frac{1}{6}\right)\left(12\lambda + 6y_t^2 - \frac{9}{2}g_2^2 - \frac{3}{2}g_1^2\right)\,, \tag{7.0.4}$$

where y_t, g_1 and g_2 are the top Yukawa, U(1) and SU(2) couplings respectively. Then, given that these couplings are of order 0.1 near the Planck scale and for a $\xi_0 \sim 10$, β_ξ

[2]See the Particle Data Group 2018 review [433] for a detailed description of the best fit values and their uncertainty. Note in particular that the top mass still has a large statistical and systematic uncertainty, and that LHC measurements have lowered the value of m_t with respect to the Tevatron results.

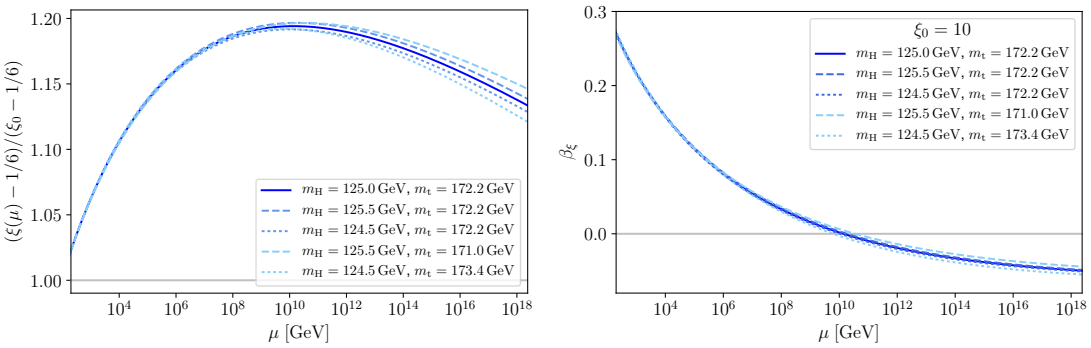

Figure 7.2. Running of the non-minimal coupling $\xi(\mu)$ and its beta function $\beta_\xi(\mu)$ following the renormalization group equations of the SM at 2-loops up to the Planck scale. The beta function is less sensitive to the values of $m_{\rm Higgs}$ and $m_{\rm top}$.

would be of order 0.01. This can be confirmed in Fig. 7.2 where we plot the 2-loop running of ξ and its beta function β_ξ using the equations of [432].

However, there is an important remark. The SM itself is renormalizable but, when we coupled it to gravity, it becomes non-renormalizable. Therefore, this scenario should be considered as an effective field theory valid up to a given cutoff scale. Near the cutoff scale, threshold corrections from integrating out heavy degrees of freedom might introduce sizable corrections. Without a specific UV completion, the non-renormalizabilty of the theory translates in the non-unique matching of low and high-energy physics [440]. In Fig. 7.1 and 7.2, the running of λ and ξ are plotted assuming that the RGE are valid up to the Planck scale without any corrections. On the one hand, the running of the SM couplings can be computed perturbatively (as we have done in Fig. 7.1 and 7.2) up to $M_{\rm pl}/\xi$. On the other hand, for sufficiently large field values, $h \gg M_{\rm pl}/\sqrt{\xi}$, the theory becomes approximately scale invariant making a perturbative treatment also possible with control of the higher order operators and a running dictated by the chiral SM. It is precisely in this region where the universal predictions of Higgs inflation take place. Conversely, in the intermediate region $M_{\rm pl}/\xi \lesssim h \lesssim M_{\rm pl}/\sqrt{\xi}$, threshold corrections generically modify the running, and in order to connect the limiting regimes one has to choose a given renormalization scheme [440]. One should note that for large non-minimal couplings $\xi \sim 10^4$ there is sizable hierarchy between the three regimes, but, for small values $\xi \sim 10$, this hierarchy is diluted. Moreover, if ξ is small, threshold corrections are relevant close to the Planck scale possibly modifying the inflationary potential. Computing the effects of the threshold corrections has become an active research area in the past years [440–442]. Here we will not attempt to compute the threshold corrections, but rather take a phenomenological approach and parametrize the running of λ and ξ at inflationary scales.

7.1 The model

In the following, we will study a critical Higgs inflation scenario in which we take into account both the running of λ and ξ. Instead of working with the full SM RGE and threshold corrections, we will take a phenomenological approach and expand the couplings around the critical point,

$$\lambda(\phi) = \lambda_0 + b_\lambda \ln^2(\phi/\mu) \,, \tag{7.1.1}$$

$$\xi(\phi) = \xi_0 + b_\xi \ln(\phi/\mu) \,. \tag{7.1.2}$$

In this context, we extend the analysis of Higgs inflation at the critical point [437, 438] by taking into account also the running of the ξ coupling through b_ξ. This extra feature in the inflationary potential can induce a brief second plateau at scales much smaller than those of the CMB, giving rise to a peak in the primordial power spectrum. Note that this approximation that we propose works very well for the inflationary potential as it has been recently shown explicitly comparing it with RGE in Ref. [443].

The action of the Higgs-inflaton with a running Higgs self-coupling and non-minimal coupling to gravity is given by

$$S = \int d^4x\sqrt{g}\left[\left(\frac{1}{2\kappa^2} + \frac{\xi(\phi)}{2}\phi^2\right)R - \frac{1}{2}(\partial\phi)^2 - \frac{1}{4}\lambda(\phi)\phi^4\right] \,, \tag{7.1.3}$$

where $\kappa^2 \equiv 8\pi G = 1/M_{\rm Pl}^2$, and we have expanded the running of the couplings around the critical point, $\phi = \mu$, as equations (7.1.1-7.1.2). After a standard metric and scalar field redefinitions,

$$g_{\mu\nu} \to \left(1 + \xi(\phi)\phi^2\right) g_{\mu\nu} \,, \tag{7.1.4}$$

$$\phi \to \varphi = \int d\phi \frac{\sqrt{1 + \phi^2(\xi(\phi) + 6(\xi(\phi) + \phi\xi(\phi)'/2)^2)}}{1 + \xi(\phi)\phi^2} \tag{7.1.5}$$

the effective inflationary potential becomes

$$V(x) = \frac{V_0\,(1 + a\,\ln^2 x)\,x^4}{(1 + c\,(1 + b\,\ln x)\,x^2)^2} \,, \tag{7.1.6}$$

with $V_0 = \lambda_0\mu^4/4$, $a = b_\lambda/\lambda_0$, $b = b_\xi/\xi_0$ and $c = \xi_0\,\kappa^2\mu^2$. The potential has a flat plateau at large values of the field $x = \phi/\mu$, see top-left panel of Fig. 7.3. Interestingly, the amplitude of the plateau is set by

$$V(\phi_{\rm CMB}) \sim b_\lambda/b_\xi^2 \,. \tag{7.1.7}$$

Thus, the small value of $H_{\rm inf}^2 = \kappa^2 V(\phi_{\rm CMB})/3 \ll M_{\rm P}^2$ is determined in this model by the RGE running of the SM Higgs couplings λ and ξ solely. This differs from the asymptotic behavior of Higgs inflation (7.0.1) and Higgs inflation at the critical point (7.0.3).

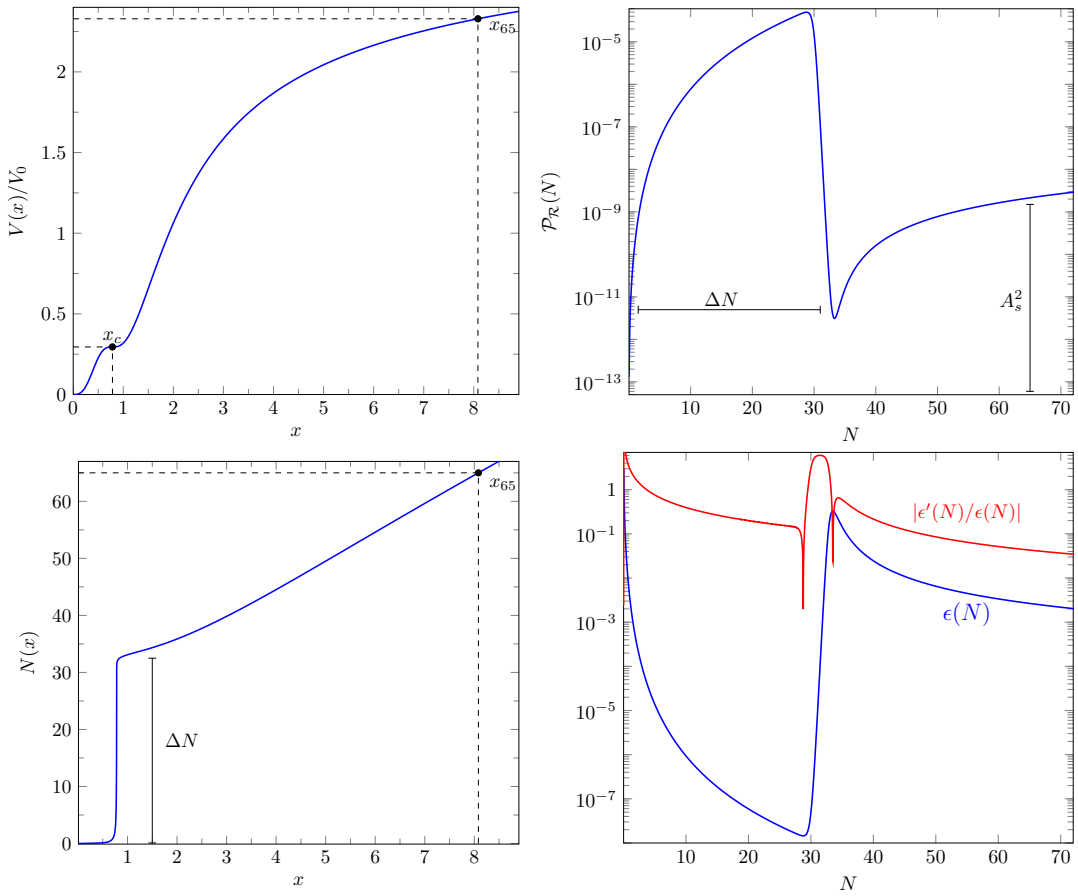

Figure 7.3. *Top panels:* the Critical Higgs Inflation potential (left) and its curvature power spectrum $\mathcal{P}_\mathcal{R}(N)$ (right). The large and broad *half-dome* peak at small scales ($N < \Delta N$) is responsible for PBH production over a wide range of masses. *Bottom panels:* evolution of the number of e-folds (left) and the slow-roll parameters (right) for the exact equations of motion.

The potential also has a short secondary plateau around the critical point, where the inflaton-Higgs slows down and induces a large peak in the curvature power spectrum. This second plateau is induced by a near-inflection point at $x = x_c$, where $V'(x_c) \simeq 0$, $V''(x_c) \simeq 0$. As a consequence, the number of e-folds has a sharp jump ΔN at that point, cf. bottom-left panel of Fig. 7.3, plus a slow rise towards larger field values, corresponding to CMB scales. This behavior is analogous to the toy model considered in section 6.2.

We can compute the inflationary evolution solving numerically equation (6.2.2). Again, one should notice that, although the slow-roll parameter $\epsilon(N) = \kappa^2\varphi'(N)^2/2$ is always smaller than one, its variation $\epsilon'(N)/\epsilon(N)$ can be larger around the near inflection point. Still, for a large set of the CHI parameter space, the inflaton slows down around x_c, producing a broad peak in the spectrum, but keeps enough inertia to cross the near-inflection point and continue rolling down the potential towards the end of in-

flation in just a few e-folds. This is exemplified in the bottom-right panel of Fig. 7.3. Thus, CHI can produce a successful inflation with a characteristic *half-dome* peak in the spectrum at small scales.

7.2 Predictions

We chose to explore the predictions of the model in terms of the height and width of the peak in the power spectrum, see top-right panel of Fig. 7.3, because these are the quantities that relate more directly to the abundance of PBHs. The height of the peak relative to the amplitude at CMB scales (A_s^2) is controlled by the closeness of x_c to a true inflection point, $V'(x_c) = V''(x_c) = 0$. The width of the peak is determined by the jump in the number of e-folds, ΔN. From the conditions in the potential, we find that there will be a true inflection point at x_c if

$$a(x_c,\, c) = \frac{4}{1 + c\, x_c^2 + 2\ln x_c - 4\ln^2 x_c}\,, \tag{7.2.1}$$

$$b(x_c,\, c) = \frac{2(1 + c\, x_c^2 + 4\ln x_c + 2c\, x_c^2 \ln x_c)}{c\, x_c^2(1 + c\, x_c^2 + 2\ln x_c - 4\ln^2 x_c)}\,. \tag{7.2.2}$$

Thus, a near-inflection point can be characterized by

$$a \to a(x_c,\, c) \quad \text{and} \quad b \to (1-\beta)\, b(x_c,\, c)\,, \tag{7.2.3}$$

where β controls the distance in parameter space to the true inflection point.[3] Then, the relative height of the peak will be inversely proportional to β and will increase with the width ΔN. We explore the (β, ξ_0, x_c, c) parameter space searching for power spectra consistent with the latest CMB constraints and producing a sizeable peak at x_c. The value of λ_0 is chosen to match to the observed CMB amplitude A_s^2. Therefore, there is a one-to-one correspondence between each point of the viable parameter space, and the parameters of the potential via Eqs. (7.2.1−7.2.2).

7.2.1 Early Universe observables

We have studied the main CMB observables (the scalar spectral index n_s, its running, $\alpha_s = dn_s/d\ln k$, and the tensor-to-scalar ratio r), as a function of (x_c, c), for different heights and widths. We find that, for each pair (β, ΔN), there are many choices of (x_c, c) that give rise to viable cosmologies. In particular, we have chosen as reference point in parameter space,

$$\beta = 10^{-5}\,,\;\; \Delta N = 33.5\,,\;\; x_c = 0.784\,,\;\; c = 0.77\,, \tag{7.2.4}$$

which give the CMB parameters

$$n_s = 0.952\,,\; r = 0.043\,,\; \alpha_s = -0.0017\,, \tag{7.2.5}$$

[3] One should notice that this parametrization is not unique but turned out to be the most practical. Other choices might allow to explore different sectors of the parameter space.

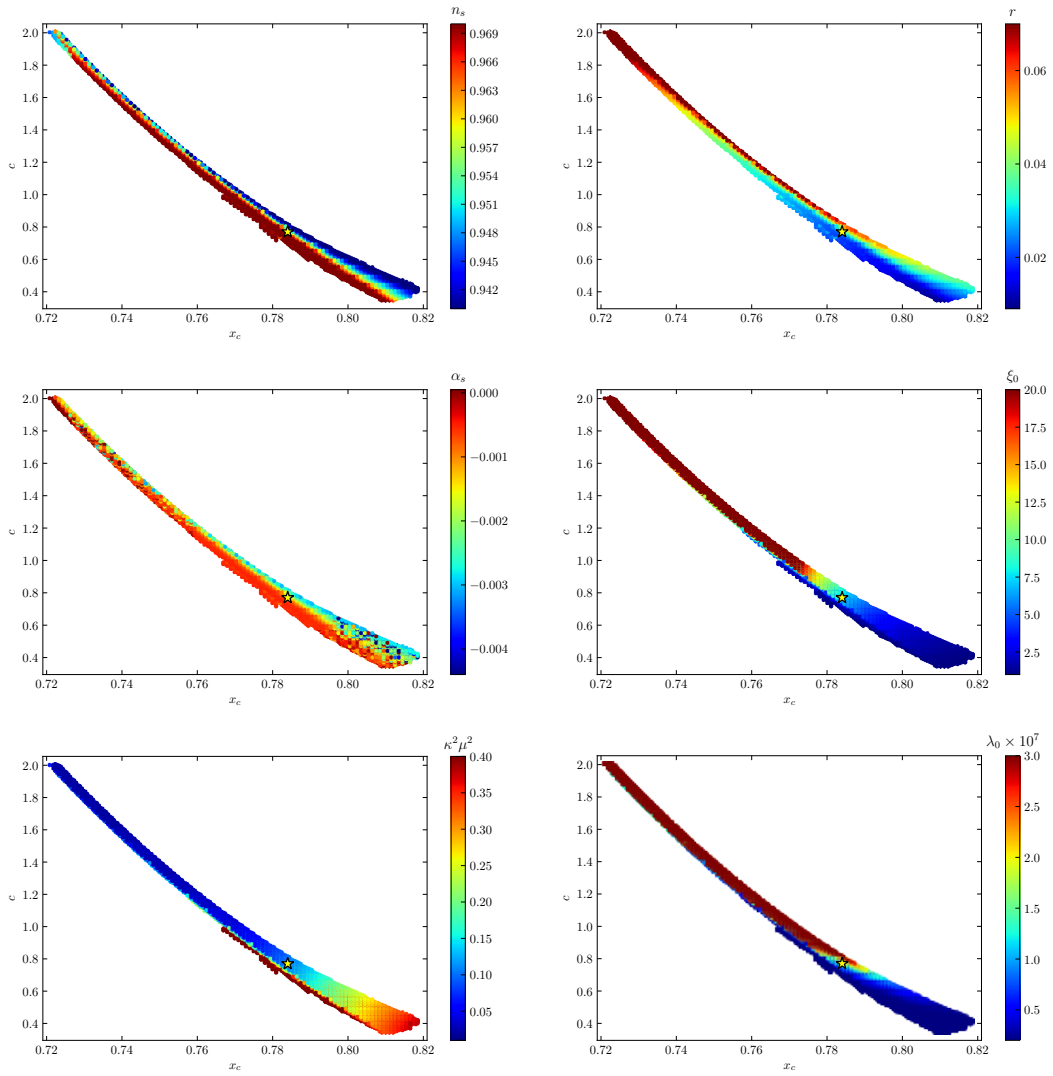

Figure 7.4. Spectral index n_s (top-left), tensor-to-scalar ratio r (top-right), running of the spectral index α_s (middle-left), non-minimal coupling ξ_0 (middle-right), scale of critical point $\kappa^2\mu^2$ (bottom-left) and λ_0 (bottom-right) in the plane of the parameter space (x_c, c). The star corresponds to the reference parameter choice $x_c = 0.784$ and $c = 0.77$ (see (7.2.4)) that yields to $n_s = 0.952$, $r = 0.043$, $\alpha_s = -0.0017$, $\xi_0 = 7.55$, $\kappa^2\mu^2 = 0.102$ and $\lambda_0 = 2.23 \times 10^{-7}$.

within the 2σ limits of Planck 2018 [444]. We plot all the other viable choices in the (x_c,c)-plane of the parameter space in Fig. 7.4, showing the values of the spectral index n_s, tensor-to-scalar ratio r and running of the spectral index α_s in the first three panels starting from the top.

We present in Fig. 7.5 the predictions of the model for a range of parameters in the (n_s, r)-plane for $\beta \in (0.1 - 9) \times 10^{-4}$, and $\Delta N \in (10, 45)$, together with the 1 and 2σ constraints from CMB anisotropies, as measured by Planck 2015, shown by the

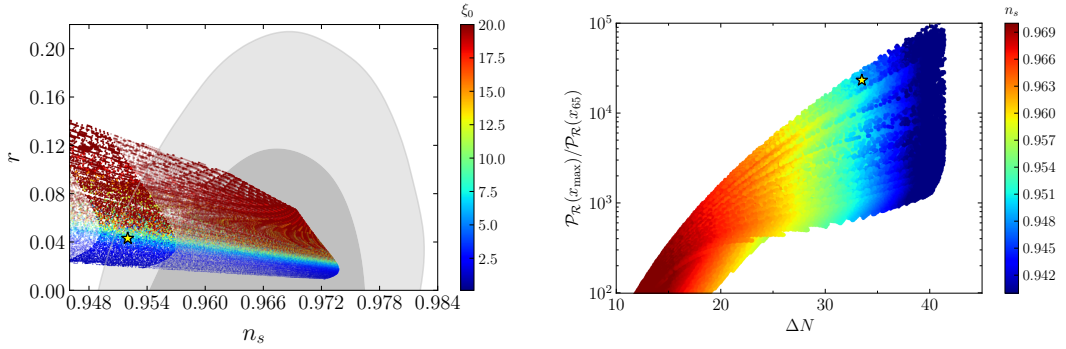

Figure 7.5. *Left-panel:* $(n_s,\, r)$-plane of CHI. The region with denser color corresponds to $\Delta N \in (30, 35)$ and the contours represents the 1 and 2σ Planck 2015 constraints for models with variable n_s, $dn_s/d\ln k$ and r, obtained from the Planck Legacy Archive. *Right panel:* height of the peak as a function of its width. In both cases, the star corresponds to the reference parameter choice with $n_s = 0.952$, $r = 0.043$, $\Delta N = 33.5$ and $\mathcal{P}_\mathcal{R}(x_{\rm max})/\mathcal{P}_\mathcal{R}(x_{65}) = 2.3 \times 10^4$ and the other points are all within $\beta \in (0.1 - 9) \times 10^{-4}$ and $\Delta N \in (10, 45)$.

grey contours.[4] We show in color the values of the non-minimal coupling ξ_0 in the $(n_s,\, r)$-plane. The region with more intense color represents cases with $\Delta N \in (30, 35)$, which produce a sufficiently large peak in the power spectrum at small scales to later give rise to PBH through gravitational collapse upon reentry. This region tends to give low spectral index, $n_s < 0.956$, and large tensor-to-scalar ratios, $r > 0.019$, while cases with lower ΔN display a better fit to Planck data but cannot generate significant populations of PBHs. In the right panel of Fig. 7.5 we show the ratio $\mathcal{P}_\mathcal{R}(x_{\rm max})/\mathcal{P}_\mathcal{R}(x_{65})$ of the amplitude of the fluctuations at its maximum, $x_{\rm max}$, over the amplitude at the inflationary plateau, x_{65}, as a function of ΔN. The color code indicates the spectral tilt n_s for each particular case. This figure shows that significantly large ratios can only be obtained for large values of ΔN, which are also associated with lower values of n_s. Although the $\Delta N \in (30, 35)$ sector of the parameter space is consistent with CMB data, having a mechanism to enlarge the ratio $\mathcal{P}_\mathcal{R}(x_{\rm max})/\mathcal{P}_\mathcal{R}(x_{65})$ could relax the constraints in the parameter space associated to Planck data. In the next chapter, we will discuss how quantum diffusion effects could enhance the peak in the power spectrum and the overall PBH production.

The reference point (7.2.4) corresponds to the model parameters

$$\begin{aligned} &\lambda_0 = 2.23 \times 10^{-7}\,,\ \xi_0 = 7.55\,,\quad \kappa^2\mu^2 = 0.102\,, \\ &b_\lambda = 1.2 \times 10^{-6}\,,\quad b_\xi = 11.5\,. \end{aligned} \tag{7.2.6}$$

In order to have a large PBH production and a good agreement with the CMB con-

[4]Note that Planck 2018 contours do not differ qualitatively from Planck 2015 results in which the original work [4] was based.

straints, the allowed range of CHI parameters can be enlarged to

$$\begin{aligned} &\lambda_0 \sim (0.01 - 8) \times 10^{-7}\,,\ \xi_0 \sim (0.5 - 15)\,, \quad \kappa^2\mu^2 \sim (0.05 - 1.2)\,, \\ &b_\lambda \sim (0.008 - 4) \times 10^{-6}\,, \quad b_\xi \sim (1 - 18)\,, \end{aligned} \tag{7.2.7}$$

for $\Delta N \in (30, 35)$. We plot the values of the non-minimal coupling ξ_0, the scale of the critical point $\kappa^2\mu^2$ and λ_0 in the (x_c, c)-plane in the last three panels of Fig. 7.4.

The question of whether these values, corresponding to the model parameters at the critical scale μ, are consistent with the values of the SM parameters at the EW scale is complicated. This is caused by the uncertainty in the matching derived from the non-renormalizability of the theory that we have discussed before. Given the latest values of $m_{\rm top}$ that we quote in (7.0.2), the values of λ_0 and b_λ that we consider for the Higgs quartic coupling are consistent with the existence of a critical point $\beta_\lambda(\mu_c) = \lambda(\mu_c) = 0$ around scales $\mu_c \sim 10^{17} - 10^{18}$ GeV as plotted in Fig. 7.1. A recent analysis claims however that the range of λ that we report in (7.2.7) is smaller by an order of magnitude than the one expected in the SM RGE (without threshold corrections) up to the Planck scale [445]. On the other hand, the non-minimal coupling of the Higgs to gravity ξ_0 is still largely unbounded (only very high values could affect particle physics experiments) and there is no prediction from the SM itself. Nonetheless, its running β_ξ is determined by the SM couplings as it can be observed for instance in the 1-loop expression (7.0.4). Another recent study, which does include the effect of threshold corrections in the running of λ, has pointed out that a value of $b_\xi \sim 10$ is larger than the one expected in the SM [442]. Following the philosophy of Higgs inflation, this analysis has a bottom-up approach, in which starting from the SM non-minimally coupled to gravity, they only introduce those higher order operators necessary for the quantum theory to remain consistent. It would be interesting to investigate if other types of higher-order operators could effectively enlarge the value of b_ξ, for instance R^2. Additionally, it would be worth exploring if a more efficient PBH production mechanism (e.g. due to quantum diffusion) could serve to lower the value of b_ξ. We leave this analysis for future work.

With all these considerations, if this CHI scenario was correct and the BBH mergers detected by LIGO were of primordial origin, future measurements of the BBH mass spectrum would allow to obtain complementary information about the SM couplings of the Higgs at high energy scales. This enables to have a large lever arm for the RGE running of these couplings from the EW scale to almost the Planck scale. Although this possibility is tantalizing, we emphasize again that a detailed analysis of the compatibility of these coefficients with the predictions of the SM non-minimally coupled to gravity, possibly with the inclusion of threshold corrections, requires further work. Moreover, as we will discuss next, the connection between GW observables and high-energy physics has also some complications.

In any case, it is also interesting to note that this CHI scenario predicts an amplitude of tensor modes that lies within the target range of present and next-generation B-mode experiments. For instance, CMB Stage-4 targets at $r < 0.002$ at 95% C.L. [261], which is an order of magnitude better than the reference value of this model $r \sim 0.04$. Moreover,

the large amplitude of curvature fluctuations a few e-folds before the end of inflation, see Fig. 7.3, may induce a significantly inhomogeneous reheating upon reentry, which could have important consequences for the reheating temperature and possibly also for the production of PBH and gravitational waves at preheating, see e.g. [446]. In particular, we find that the energy density at the end of inflation is $\rho_{\rm end} = 4 \times 10^{63}$ GeV^4 and the estimated reheating temperature (for $g_* = 106.75$), $T_{\rm rh} = 3.2 \times 10^{15}$ GeV, is relatively high, justifying our choice of $N = 65$ e-folds of inflation.

7.2.2 Late Universe observables

We now move to possible observables of this CHI scenario in the late universe. We use the Press-Schechter formalism of gravitational collapse to compute the probability that a given horizon-sized volume forms a PBH when a large curvature fluctuation, $\zeta > \zeta_c$, reenters the horizon during the radiation era, and not even radiation pressure can prevent collapse. Under the Gaussian approximation, the fraction of PBHs at formation can be computed from (6.0.4). The mass of the PBH at formation is essentially given (within an order-one efficiency factor γ) by the total mass within the horizon at the time of reentry, see eq. (6.0.2). In our case, for the large and wide peak in $\mathcal{P}_\zeta(k)$ at small scales, one finds an approximate lognormal distribution of masses for PBH,

$$P(M) = \frac{A\,\mu}{M\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\ln^2(M/\mu)}{2\sigma^2}\right), \tag{7.2.8}$$

with a sharp drop at high masses due to the half-dome shape of the peak, see Fig. 7.3. This characteristic shape also shifts the peak of the mass spectrum to higher values since the PBH mass exponentially depends on the number of e-folds at reentry.

The distribution of PBHs at formation is characterized by the physics of inflation through $\mathcal{P}_\zeta(k)$ and by the conditions at the time of formation through ζ_c. As we have pointed out in chapter 6, the emphasis of our work is in the early universe physics and not the formation mechanism itself. Accordingly, we will consider the formation of PBHs in CHI for a fixed value of ζ_c. At the end, refining the choice of ζ_c or changing the criterion of formation accounts for searching a different region of the parameter space. Since we will see in the next chapter 8 that there are indeed corrections to the initial assumptions on the primordial spectrum (induced by quantum diffusion effects), we will limit to this methodology. We stress that our goal is to show that this class of CHI models have the potential to produce large amounts of PBHs rather than pin-pointing a particular scenario.

From the time of formation to the matter-radiation equality, the fraction of PBHs linearly increases with the scale factor

$$\beta_{\rm eq}(M) = \beta_f(M) \cdot a_{\rm eq}/a(t_{_M}), \tag{7.2.9}$$

since the radiation dilutes faster than matter with the expansion. During this period there is also evaporation of the lightest PBHs due to Hawking radiation. We find that

for the range of $\Delta N \in (30-35)$, PBHs can constitute the total DM at equality, i.e.

$$\Omega_{\rm PBH}^{\rm eq} = \int \beta_{\rm eq}(M)\, d\ln M = 0.42\,, \tag{7.2.10}$$

within the uncertainty range of $\zeta_c \sim (0.05-1)$ [403]. For the reference point in parameter space that we have chosen, we use $\zeta_c = 0.052$ and $\gamma = 0.4$. Here we do not consider any non-linear growth in mass before equality, which might increase the abundance of PBHs at equality $\Omega_{\rm PBH}^{\rm eq}$.

From equality to present times, the mass distribution will shift to higher masses due to merging and accretion. Key aspects in this process are the spatial distribution of PBHs and the time in which the BH binaries form. Binaries could form during radiation epoch since the mean distance of PBHs compare to the Hubble horizon decreases with the expansion [238]. They could also form after equality due to close encounters of PBHs in dense environments [240,241]. As any other non-relativistic matter, PBHs will cluster at late times due to gravity. However, the question on the initial clustering has been subject of intense debate recently, where the highly cluster assumption [447] has been challenged by [448] advocating to a Poisson distribution, which has also been disputed by [449]. Knowing the initial spatial distribution is relevant for instance in order to properly asses the number of binaries formed in the early universe and the mass growth in that period, as well as to evade certain constraints [450]. In the remaining of the chapter, we will restrict to the highly cluster assumption used in [4].

In this CHI scenario, there is a very wide peak in the matter spectrum at small scales. In order to exactly determine the mass distribution of PBHs today, one would have to solve the non-linear evolution with a N-body simulation. Following the highly clustering scenario of Ref. [447], we roughly estimate the growth in PBH masses by a factor 3×10^7. In this case, we find that the peak of the lognormal distribution corresponds today to approximately $\mu_{\rm PBH} \simeq 11\,M_\odot$ and the lognormal dispersion to $\sigma_{\rm PBH} \simeq 0.8$. Note that the mean of the PBH distribution is determined by the location of the maximum of the power spectrum $N_{\rm peak}$, leading to

$$\mu_{\rm PBH} \sim 10\,M_\odot \cdot e^{2(N_{\rm peak}-28.8)}\,, \tag{7.2.11}$$

while the variance is more sensitive to the width of the peak ΔN. For the range we are considering, $\Delta N \in (30-35)$, then $\sigma_{\rm PBH} \sim (0.6-1)$ and $N_{\rm peak} \sim (25-30)$. Therefore, DM would be dominated today by PBH with masses in the range from 0.01 to $100\,M_\odot$. As a consequence, under the mentioned assumptions, the CHI scenario is able to generate the high-mass BBH mergers that have been observed by LIGO.

PBH scenarios like CHI can be probed with GWs by several means. A clear smoking gun in favor of the primordial origin would be to detect GWs from a binary in which one of the BHs has a sub-solar mass. This is because there is no known stellar evolution channel to form them due to the Chandrasekhar limit [451]. LIGO has performed specific searches in their first [452] and second [453] observing runs, with no viable candidate found. However, one should note that these searches use template banks for binaries with total mass in the range $0.4-4.0M_\odot$, not accounting for small mass ratios $q = m_2/m_1 \ll 1$.

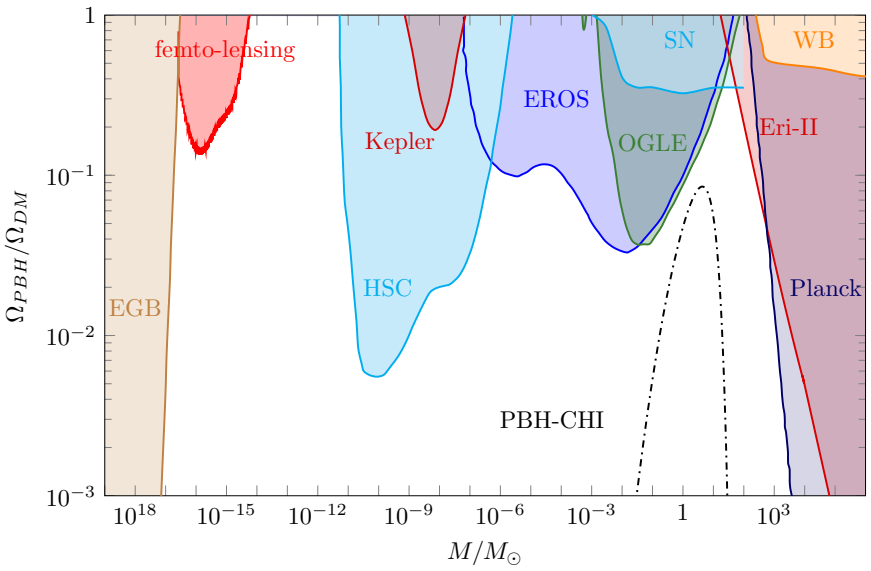

Figure 7.6. Constraints on monochromatic distributions of PBHs from Extragalactic Gamma Background (EGB) [248], femto-lensing of gamma ray bursts [233], micro-lensing (HSC [231], Kepler [232], EROS [229] and OGLE [230]), supernova lensing (SN) [235], Wide Binaries (WB) [236], Eridanus II (Eri-II) [237] and the CMB (Planck) [247]. For details on each constraint we refer to the discussion in section 1.2.2. The extended mass distribution of PBHs produced in the reference Critical Higgs Inflation model (PBH-CHI, dashed-dotted line) peaks at $\mu_{\rm PBH} \simeq 11\, M_\odot$ with dispersion $\sigma_{\rm PBH} \simeq 0.8$. We discuss the constraints on this scenario at the end of section 7.2.2.

Other possibilities to probe the PBH hypothesis is to analyze the mass spectrum, spin distribution, eccentricities and merger rate as a function of redshift of the observed BBHs. From them, probably the most relevant in the near future would be the effective spin. Since PBH are formed during the radiation dominated era form the collapse of large curvature perturbations, they are expected to be form with no spin. Thus, the observed BBHs should have a spin distribution peaking at 0. Interestingly, this picture is compatible with present O1-O2 data as it can be observed in Fig. 1.4 (see [454] for a detailed analysis). In the more distant future, observing large merger rates at high redshift with third generation GW detectors could be an indicator of the formation of the BHs in the early universe.

Apart from the direct GW emission from BBH mergers that LIGO is sensitive to, there are several stochastic backgrounds coming from different epochs that could be detected. One GW background comes from unresolved BBH mergers since equality, with a power law spectrum,

$$h^2\,\Omega_{\rm GW}(f) = 8\times 10^{-15}\;\tau_m\, f^{2/3}({\rm Hz})\,\mu^{5/3}(M_\odot)\,R(\sigma)\,,$$

where $\tau_m \sim 50$ events/yr/Gpc^3 is the BHB merger rate and $R(\sigma)$ is an exponentially growing function of σ, see [455]. In the near future we may be able to detect this irreducible GW background with LISA [456]. A totally different background arises from

second-order anisotropic stresses induced by large curvature fluctuations at horizon reentry when PBH formed, which depending on the time of re-entry could peak from sub nHz frequencies relevant for SKA to 1-100 mHz frequencies relevant for LISA, BBO or DECIGO [457]. In fact a recent analysis showed the possible detectability of this second-order background in CHI [458].

Finally, it is relevant to compare the predicted distribution of PBHs against present constraints. PBH constraints differ significantly in the mass window and the epoch or physics they probe (see discussion in 1.2.2). Generically, PBHs can be constrained due to the energy injections to the primordial plasma when they accrete matter before recombination and due to the Hawking radiation they would emit today if they were evaporating. They can also be probed using lensing techniques that can distinguish between the fluid and the compact nature of DM. Dynamical effects are another source of constraints as well as the already mentioned GW observations. For all of them, it is important to remember that linking the astrophysical or cosmological observations with constraints on particular PBH scenarios requires of the same set of assumptions that are necessary to move from a theoretical model to its predictions. Moreover, different constraints apply to different scales and periods of the cosmic history, which affects for instance to PBH distributions that evolve with time. This means that constraints, as theoretical predictions, should be considered with certain caution and, when possible, referred with their given assumptions.

In Fig. 7.6 we summarize the state-of-the-art constraints on the fraction of the DM composed of PBHs given by the ratio $\Omega_{\rm PBH}/\Omega_{\rm DM}$.[5] The constraints plotted apply to monochromatic distributions in which all the PBHs have the same mass. We estimate the effect of the non-zero width of the distribution as in Ref. [459]. Since in our case the width is not very large, $\sigma_{\rm PBH} \simeq 0.8$, the constraints do not significantly change w.r.t. the monochromatic case. We include also the expected PBH distribution today in CHI under the highly clustered assumption. Taken at face value, the microlensing and SN lensing constraints would imply that the fraction of the DM in the form of PBHs in the mass range $1 - 100 M_\odot$ could be no more than $\mathcal{O}(0.1)$. However, it is important to emphasis that the constraints in Fig. 7.6 do not take into account the spatial distribution of PBHs and consider them as uniformly distributed on space. This affects for instance microlensing constraints, since the probability to find a PBH in the line of sight of the source is reduced [450]. SN lensing constraints might be as well affected in the case of clustered PBHs as argued in [460]. Specific reanalysis of the data considering different spatial distributions should be performed in the future along with a more precise theoretical modeling.

[5]Note that they have changed for instance compared to the ones we presented in [4]. Most notably, the reanalysis of HSC data taking into account the wave-effects has opened a window for small PBH in the range $10^{-14} - 10^{-11} M_\odot$ [231].

CHAPTER

8

Effects of quantum diffusion

Non-perturbative quantum corrections during inflation become relevant when the curvature perturbations are large, approaching $\zeta \sim 1$, so that the perturbative expansion breaks down. Since PBH production relies precisely on large primordial perturbations produced during inflation (see chapter 6), it is natural to explore the back-reaction of these quantum fluctuations on the classical inflationary dynamics, to asses if they affect the production of PBHs. This will be the main objective of this chapter. For this goal, stochastic inflation [461] is the appropriate tool since it is an effective theory that tracks stochastic effects of the short wavelength modes on the long wavelength ones outside the horizon. Within the SR approximation, the stochastic evolution is determined by the standard Langevin equation for the coarse-grained field $\bar{\phi}$,

$$\frac{d\bar{\phi}}{dN} = -\frac{V'}{3H^2} + \frac{H^2}{4\pi}\xi_\phi(N)\,, \tag{8.0.1}$$

where $\xi_\phi(N)$ is a Gaussian white noise, meaning $\langle \xi_\phi(N)\rangle = 0$ and $\langle \xi_\phi(N)\xi_\phi(N')\rangle = \delta(N-N')$. From this expression we can already anticipate that quantum diffusion effects will be relevant whenever *i)* the drift term $V'/3H^2$ is small, i.e. when the field slows down $\epsilon \ll 1$, or *ii)* the noise becomes large. We have seen that in the quasi-inflection point scenarios there are both a region with $\epsilon \ll 1$ and with large quantum fluctuations, suggesting that quantum diffusion effects can be relevant.

Because in this class of models there are deviations from SR, we have to consider the stochastic evolution of both the coarse-grained field $\bar{\phi}$ and its canonical momentum $\bar{\pi}_\phi$ with their respective noises ξ_ϕ and ξ_π. Building on the Hamilton-Jacobi formalism for stochastic inflation [462, 463] summarized in Sec. 8.1, we develop a methodology in Sec. 8.2 to obtain the PDF of the stochastic fluctuations from their correlation functions solving a system of first-order differential equations. This method allows to solve iteratively the statistical moments of the PDF up to a given order. We can then compute analytically [464] the corrections of this non-Gaussian (NG) contributions to the probability to form a PBH.

Higher order correlators can be very significant because the fraction of PBHs formed is very sensitive to the tails of the PDF. Several studies have discussed the importance of considering NG corrections to PBH formation [424, 465–470], as well as the relevance of quantum diffusion in the context of hybrid inflation [224, 399] and single-field models [471–475]. Here, we place a full analysis of the quantum diffusion beyond SR, computing

the noise and the non-Gaussian corrections. We find that quantum diffusion can significantly alter the classical prediction. Our results have significant implications for the abundance of PBHs that we discuss in Sec. 8.3. These findings extend to several classes of inflationary models. In particular, we expect them to be relevant for the Critical Higgs Inflation presented in chapter 7.

In order to gain further insight on the effects of quantum diffusion, we develop an alternative method to calculate directly the tail of the PDF in chapter 9. The advantage of this is that one then has the information from all the statistical moments at once. Although we limit to SR, we will find that quantum diffusion generically makes the tail to follow χ^2 statistics rather than Gaussian. In other words, the tail is much larger because instead of decaying as $e^{-\zeta^2/2\sigma}$, the PDF decays as $e^{-\zeta/2\sigma}$, were $\sigma^2 = \mathcal{P}_\zeta$.

8.1 Stochastic inflation beyond slow-roll

In section 6.2 we studied the classical dynamics of inflation with a quasi-inflection point. We found that the evolution exits the SR attractor leading to a phase, when $3 - \epsilon + \epsilon_2 < 0$, in which the quantum modes exponentially grow outside the horizon. We now ask if quantum corrections can modify the inflationary observables. For that purpose, stochastic inflation is the appropriate tool since it tracks the effects of the short-wavelength modes exiting the horizon on the long-wavelength ones by coarse-graining the field [461, 476]. Thus, this formalism captures stochastically quantum correction to the classical evolution, which is exactly our goal.

Since we are studying a system beyond slow-roll, we have to solve for both the field and its canonical momentum. Therefore, it is appropriate to work in phase space using the Hamilton-Jacobi formalism. The analysis of stochastic corrections to the phase space dynamics has been addressed previously in Ref. [462, 463]. Recently, a reanalysis [477] has confirmed the validity of the stochastic paradigm beyond slow-roll. In this first introduction to stochastic inflation we will follow closely the developments of Ref. [463]. We use the number of e-folds as time variable[1], $dN = Hdt$, and define the conjugate momenta $\pi_\phi \equiv d\phi/dN$. We then split the field in long and short wavelengths as

$$\phi = \bar{\phi} + \phi_s \tag{8.1.1}$$

$$\pi_\phi = \bar{\pi}_\phi + (\bar{\pi}_\phi)_s \,, \tag{8.1.2}$$

where quantities with an over bar represent the coarse-grained field and the subindex s indicates the short-wavelength modes. The evolution of the long wavelengths can be obtained by solving the Langevin equations

$$\frac{d\bar{\phi}}{dN} = \bar{\pi}_\phi + \xi_\phi \,, \tag{8.1.3}$$

$$\frac{d\bar{\pi}_\phi}{dN} = -(3-\epsilon)\left(\bar{\pi}_\phi + \kappa^{-2}(\ln V)_{,\phi}\right) + \xi_\pi \,, \tag{8.1.4}$$

[1] See a detailed discussion in Ref. [478] of why this is the appropriate variable.

where ξ_ϕ and ξ_π are the noise associated to $\bar{\phi}$ and $\bar{\pi}_\phi$ respectively. The noise terms encapsulate all the physics associated to the short-wavelength modes. They can be computed by integrating the quantum modes in Fourier space with a time dependent cut-off tracking the horizon size. We leave the detailed derivation of the noise in Appendix C.1. The important result is that the 2-point correlations of the noise can be associated to the power spectrum of the fluctuations [463]

$$\Xi_{AB}(N,N') \equiv \langle 0|\xi_A(N)\xi_B(N')|0\rangle = \frac{d\ln k_\sigma}{dN}\mathcal{P}_{AB}(k_\sigma,N)\cdot\delta(N-N')\,, \tag{8.1.5}$$

where $k_\sigma = \sigma aH$ is the cutoff (with $\sigma \ll 1$) and A, B could be either $\delta\phi$ or $\delta\pi$. We are defining the power spectrum as

$$\mathcal{P}_{AB}(k_\sigma,N) = \frac{k_\sigma^3(N)}{2\pi^2}A_{k_\sigma}(N)B^\star_{k_\sigma}(N)\,. \tag{8.1.6}$$

For instance, the noise associated to the coarse-grained field $\bar{\phi}$ is characterized by

$$\Xi_{\phi\phi}(N) = (1-\epsilon(N))\mathcal{P}_{\delta\phi}(N)\,, \tag{8.1.7}$$

since $d\ln k_\sigma/dN = 1-\epsilon_1$. One should note that in general there will be also a contribution to the noise in the momentum, $\Xi_{\pi\pi}\neq 0$, and a correlation between them, $\Xi_{\phi\pi}\neq 0$. It is only in the slow-roll approximation that $\Xi_{\pi\pi}\approx\Xi_{\phi\pi}\approx 0$. The fact that the noise is defined through its 2-point function (8.1.5) is a consequence of working at linear order in perturbation theory for the short-wave modes. Moreover, the noise is uncorrelated at different times, $\Xi_{AB}(N,N')\propto\delta(N-N')$, i.e. it is white/Markovian noise. This is a result of choosing a Heaviside window function as the momentum cutoff [462, 463] (see details in App. C.1). Alternative choices still give a correlation which is exponentially suppressed at times $\Delta N = N-N' \geqslant 1$ [479].

The power spectrum of the fluctuations can be computed by solving the linearized equations for the short-wavelength modes (in phase space)

$$\frac{d\delta\phi_k}{dN} = \delta\pi_k\,, \tag{8.1.8}$$

$$\frac{d\delta\pi_k}{dN} = -(3-\epsilon)\delta\pi_k - \left[\left(\frac{k}{aH}\right)^2 + \frac{3-\epsilon}{\kappa^2}\frac{V_{,\phi\phi}}{V} - 2\epsilon(3-\epsilon+\epsilon_2)\right]\delta\phi_k\,. \tag{8.1.9}$$

Note that we have included in the effective mass term a piece $-2\epsilon(3-\epsilon+\epsilon_2)$ that accounts for the effect of the metric fluctuations and, thus, the curvature perturbation obeys $\zeta_k = \kappa\delta\phi_k/\sqrt{2\epsilon}$. Then, the spectrum derived from the above evolution equations is equivalent to the Mukhanov-Sasaki formalism[2]. The noise associated to $\bar{\phi}$ in our inflationary model with a quasi-inflection point is plotted in Fig. 8.1. Importantly, the exponential amplification of the curvature fluctuations in the regime beyond SR with

[2]In fact, it is easy to find the equivalence between both approaches if $V_{,\phi\phi}$ is expressed in terms of slow-roll parameters ϵ using that $V = (3-\epsilon)H^2/\kappa^2$ and $d\phi/dN = \sqrt{2\epsilon}/\kappa$.

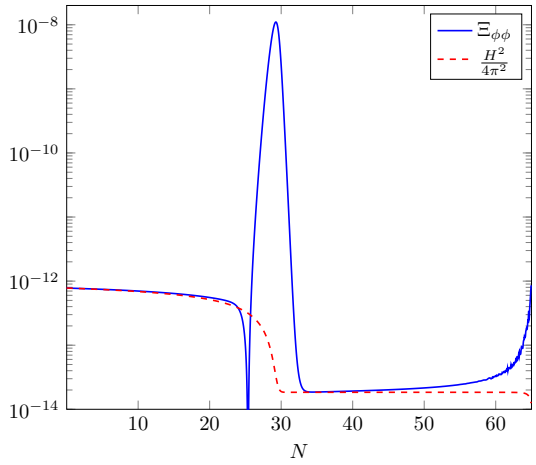

Figure 8.1. Stochastic noise $\Xi_{\phi\phi}$ as a function of the number of e-folds for the toy model (6.2.1). There is an enhancement of $\Xi_{\phi\phi}$ in the region in which $3-\epsilon+\epsilon_2<0$ with respect to the slow-roll result $H^2/4\pi^2$.

$3-\epsilon+\epsilon_2<0$ (recall Fig. 6.1) produces a peak in $\Xi_{\phi\phi}$ with respect to the usual SR result where $\Xi^{\rm SR}_{\phi\phi}=H^2/4\pi^2$.

One important remark about the evolution of the quantum fluctuations is that they depend on the dynamics of the coarse-grained fields through ϵ_n that depend on $\bar{\pi}_\phi$ and the potential V that depends on $\bar{\phi}$. Therefore, the noise depends on the coarse-grained fields too. This implies that one would need to solve simultaneously the stochastic evolution dictated by the Langevin equations with the computation of the noise, which is a rather involved numerical problem. One way to proceed would be through an iterative process in which the noise term generated by quantum fluctuations will back-react on the classical background trajectory. This new background will then be used to compute the stochastic-corrected quantum noise and thus iteratively generate the coarse-grained fields. We expect this iterative process to converge rapidly. As a first approximation, we computed the noise from the quantum fluctuations over the classical trajectory and then solved the stochastic equations. We have checked that the correction due to the first iteration is already subdominant, justifying this approximation. In the next section we will detail this computation.

8.2 Correlation functions

From the system of Langevin equations for the coarse-grained fields $\bar{\phi}$ and $\bar{\pi}_\phi$ we can obtain the associated Fokker-Planck equation[3] for the probability density function (PDF) $P(N;\mathbf{\Phi})$

$$\frac{dP(N;\mathbf{\Phi})}{dN}=-\frac{\partial}{\partial\Phi_A}\left(D_A\cdot P(N;\mathbf{\Phi})-\frac{\Xi_{AB}}{2}\frac{\partial P(N;\mathbf{\Phi})}{\partial\Phi_B}\right), \tag{8.2.1}$$

[3]One should remember that the connection between the Langevin and Fokker-Planck equations exists because the noise Ξ_{AB} is white and Gaussian (see for instance a derivation in App. B of [462]).

where $\mathbf{\Phi}$ is the vector field of $\bar{\phi}$ and $\bar{\pi}_\phi$, and the the indices A, B sum over $\bar{\phi}$ and $\bar{\pi}_\phi$. The vector $\mathbf{D}$ represent the drift with components

$$D_\phi = \bar{\pi}_\phi \,, \tag{8.2.2}$$

$$D_\pi = -(3-\epsilon)\left(\bar{\pi}_\phi + \kappa^{-2}(\ln V)_{,\phi}\right) \,, \tag{8.2.3}$$

which follow from the Langevin equation. We will treat Ξ_{AB} as a function of time only.

We are interested in computing the effects of the noise on the classical trajectory. For that purpose, we can define the *stochastic fluctuation* as the difference of the coarse-grained field with respect to the classical trajectory

$$\delta\phi = \bar{\phi} - \phi_c \tag{8.2.4}$$

$$\delta\pi = \bar{\pi}_\phi - \pi_c \,, \tag{8.2.5}$$

and solve their evolution. Since the Fokker-Planck equation is hard to solve numerically, we can rewrite the problem in terms of the statistical moments of the PDF $\langle \delta\phi^m \delta\pi^n \rangle$. A general correlation function of the stochastic fluctuations is defined as

$$\langle \delta\phi^n \delta\pi^m \rangle(N) = \int d\bar{\pi}_\phi \int d\bar{\phi}\, (\bar{\phi} - \phi_c(N))^n (\bar{\pi}_\phi - \pi_c(N))^m P(N; \bar{\phi}, \bar{\pi}_\phi) \,. \tag{8.2.6}$$

Taking a time derivate of $\langle \delta\phi^m \delta\pi^n \rangle$ and substituting the Fokker-Planck equation (8.2.1), we obtain the general equation dictating its evolution

$$\begin{aligned}\frac{d\langle \delta\phi^n \delta\pi^m \rangle}{dN} &= n\left(\langle \delta\phi^{n-1}\delta\pi^m D_\phi \rangle - \langle \delta\phi^{n-1}\delta\pi^m \rangle D_\phi^c\right) + m\left(\langle \delta\phi^n \delta\pi^{m-1} D_\pi \rangle - \langle \delta\phi^n \delta\pi^{m-1}\rangle D_\pi^c\right) \\ &+ \frac{1}{2} n(n-1)\Xi_{\phi\phi}\langle \delta\phi^{n-2}\delta\pi^m \rangle + \frac{1}{2} m(m-1)\Xi_{\pi\pi}\langle \delta\phi^n \delta\pi^{m-2}\rangle \\ &+ \frac{1}{2} n\, m \left(\Xi_{\phi\pi} + \Xi_{\pi\phi}\right)\langle \delta\phi^{n-1}\delta\pi^{m-1}\rangle \,,\end{aligned} \tag{8.2.7}$$

where D_ϕ^c and D_π^c are the drift terms evaluated at the classical trajectory. For generic drift functions D_ϕ and D_π, we can obtain their statistical average expanding in powers of $\delta\phi$ and $\delta\pi$. In this sense, the terms D_ϕ^c and D_π^c will be killing the first term of the expansion. Another interesting point is that only the symmetric part of the noise correlation Ξ enters in the equations. In Appendix C.2 we provide the detailed form of this equation expanding the drift terms.

In order to obtain the PDF exactly, one would need to solve the infinite system of first-order differential equations described by (8.2.7). In practice, one can truncate the series at a given order in which the higher moments are subdominant. The fact that we are solving first-order differential equations speeds the numerical analysis. Moreover, truncating at a given order n, we can always solve the evolution of the statistical moments of order n which will only depend of the others of the same order. This serves to obtain the leading contribution to each order. We will discuss this in the next subsection. Later, we will comment on how to reconstruct the PDF from the n-point correlation functions.

8.2.1 Power Spectrum

The basic observable from inflation is the power spectrum of scalar perturbations. In order to obtain it, we have to compute the 2-point correlation functions. At lowest oder, their evolution is described by the following system of equations

$$\frac{d}{dN}\begin{pmatrix}\langle\delta\phi^2\rangle\\ \langle\delta\phi\delta\pi\rangle\\ \langle\delta\pi^2\rangle\end{pmatrix}=\begin{pmatrix}0 & -2 & 0\\ g(\epsilon_n) & -f(\epsilon_n) & -1\\ 0 & 2g(\epsilon_n) & -2f(\epsilon_n)\end{pmatrix}\begin{pmatrix}\langle\delta\phi^2\rangle\\ \langle\delta\phi\delta\pi\rangle\\ \langle\delta\pi^2\rangle\end{pmatrix}+\begin{pmatrix}\Xi_{\phi\phi}\\ \Xi^s_{\phi\pi}\\ \Xi_{\pi\pi}\end{pmatrix} \tag{8.2.8}$$

where we have defined

$$f(\epsilon_n)\equiv 3(1-\epsilon)+(\ln V)_\phi\pi_c = 3-\epsilon\left(1-\frac{\epsilon_2}{3-\epsilon}\right), \tag{8.2.9}$$

$$g(\epsilon_n)\equiv\frac{(3-\epsilon)}{\kappa^2}(\ln V)_{\phi\phi} = -\frac{\epsilon_2}{2}\left(f(\epsilon_n)+\frac{1}{2}\epsilon_2+\epsilon_3\right), \tag{8.2.10}$$

and $\Xi^s_{\phi\pi}$ denotes the symmetrization of the non-diagonal noise, i.e. $\Xi^s_{\phi\pi}=(\Xi_{\phi\pi}+\Xi_{\pi\phi})/2$. Once we have solved these equations, we can compute the power spectrum noting that the stochastic curvature perturbation ζ, in our gauge choice[4], follows from

$$\langle\zeta^2\rangle=\frac{\kappa^2}{2}\frac{\langle\delta\phi^2\rangle}{\epsilon}. \tag{8.2.11}$$

Consequently, we can obtain the power spectrum from

$$\mathcal{P}_\zeta=\frac{d\langle\zeta^2\rangle}{d\ln k}=\frac{1}{1-\epsilon}\frac{d\langle\zeta^2\rangle}{dN}=\frac{1}{1-\epsilon}\frac{\kappa^2}{2\epsilon}\left(\frac{d\langle\delta\phi^2\rangle}{dN}-\epsilon_2\langle\delta\phi^2\rangle\right). \tag{8.2.12}$$

If we now introduce the equations for the 2-point function, we get

$$\mathcal{P}_\zeta=\frac{1}{1-\epsilon}\frac{\kappa^2}{2\epsilon}\left(\Xi_{\phi\phi}-2\langle\delta\phi\delta\pi\rangle-\epsilon_2\langle\delta\phi^2\rangle\right). \tag{8.2.13}$$

where

$$\langle\delta\phi\delta\pi\rangle=\frac{f\,g}{f^2+g}\langle\delta\phi^2\rangle+\frac{1}{2(f^2+g)}\left(\frac{d\langle\delta\pi^2\rangle}{dN}-\Xi_{\pi\pi}-2f\left(\frac{d\langle\delta\phi\delta\pi\rangle}{dN}-\Xi_{\phi\pi}\right)\right). \tag{8.2.14}$$

Noticeably, the power spectrum does not only depend on $\langle\delta\phi^2\rangle$ and $\Xi_{\phi\phi}$. Whenever we are in the SR regime, the derivatives of $\langle\delta\phi\delta\pi\rangle$ and $\langle\delta\pi^2\rangle$ can be neglected as well as their associated noise terms. At leading order in the SR parameters, we have that $f\approx 3$ and $g\approx -3\epsilon_2/2$. This means that $\langle\delta\phi\delta\pi\rangle\approx-\epsilon_2/2$ and thus we recover the usual expression for the power spectrum

$$\mathcal{P}_\zeta\simeq\frac{\kappa^2}{2\epsilon}\Xi_{\phi\phi}. \tag{8.2.15}$$

[4]We remind the reader that this is because we have already incorporated the curvature fluctuations in the effective, time-dependent mass of $\delta\phi$ (see discussion in Sec. 8.1).

Note that if the noise is properly computed accounting for the growth after horizon crossing as previously discussed, this power spectrum is equivalent to the standard Mukhanov-Sasaki formalism. This result confirms that, at leading order, SR stochastic inflation gives the same results as the curved space, quantum field theory techniques [478]. For completeness, we derive this result also directly from SR stochastic inflation in Appendix C.3.

However, whenever we are beyond SR, there could be additional contributions to the power spectrum. This is similar to the case of hybrid inflation, were the diffusion due to the second field can significantly amplify the power spectrum [224]. For the case of single field inflation with a quasi-inflection point, any deviation would be around this critical point. Since (8.2.8) is a system of first order differential equations with non-constant coefficients there is no generic analytical solution. We will thus solve the system numerically.

One should note that whenever the quantum fluctuations are small, the combination of $-2\langle\delta\phi\delta\pi\rangle - \epsilon_2\langle\delta\phi^2\rangle$ is subdominant with respect to $\Xi_{\phi\phi}$ and therefore, the usual expression for $\mathcal{P}_\zeta$ in SR is recovered, cf. (8.2.12). However, since in our case short wavelength fluctuations become large, introducing a peak in the noise, they also produce a peak in $\langle\delta\phi^2\rangle$ and $\langle\delta\phi\delta\pi\rangle$ and their contribution is no longer negligible. As a consequence, there will be a peak in the power spectrum produced by the diffusion. This peak will coincide with the maximum of the function multiplying $\langle\delta\phi^2\rangle$, i.e. $\frac{-2fg}{f^2+g} - \epsilon_2$. In fact, around this peak the power spectrum can be approximated by

$$\mathcal{P}_\zeta \approx \frac{\kappa^2}{2\epsilon}\left(\Xi_{\phi\phi} - c\,\epsilon_2\right) , \tag{8.2.16}$$

where c is a constant. This is because $\frac{-2fg}{f^2+g} - \epsilon_2 \approx -\epsilon_2$, $\langle\delta\phi^2\rangle \approx c$ and the contribution of the other terms in (8.2.14) is subdominant. Then, the location of the diffusion peak $N_{\rm max}$ is given by the maximum of ϵ_2/ϵ, which can be solved from

$$\epsilon_3(N_{\rm max}) = \epsilon_2(N_{\rm max})\,. \tag{8.2.17}$$

In addition, we find that the relative height of the diffusion peak compared to the standard peak in $\mathcal{P}_\zeta$ scales with the size of the peak of the noise $\Xi_{\phi\phi}^{\rm max}$. This is sensitive to the transition between the first and second plateau of the potential or, in other words, to the ratio between the maximum and the minimum of ϵ. This tells us that the more the inflaton slows down approaching the quasi-inflection point, the higher the diffusion peak will be. Importantly, this enhancement of the power spectrum by quantum diffusion emerges on top of the peak due to the classical dynamics. This implies that one obtains much higher amplitudes in the power spectrum than expected with the standard techniques.

In Fig. 8.2, we present power spectrum for the curvature fluctuations for our inflationary model with a quasi-inflection point resulting from stochastic inflation beyond SR, Eq. (8.2.12), the Mukhanov-Sasaki equation, Eq. (6.1.5), and the SR approximation, Eq. (6.1.6). In addition to the differences between the Mukhanov-Sasaki equation and the SR approximation that we discussed in section 6.1, quantum diffusion introduces

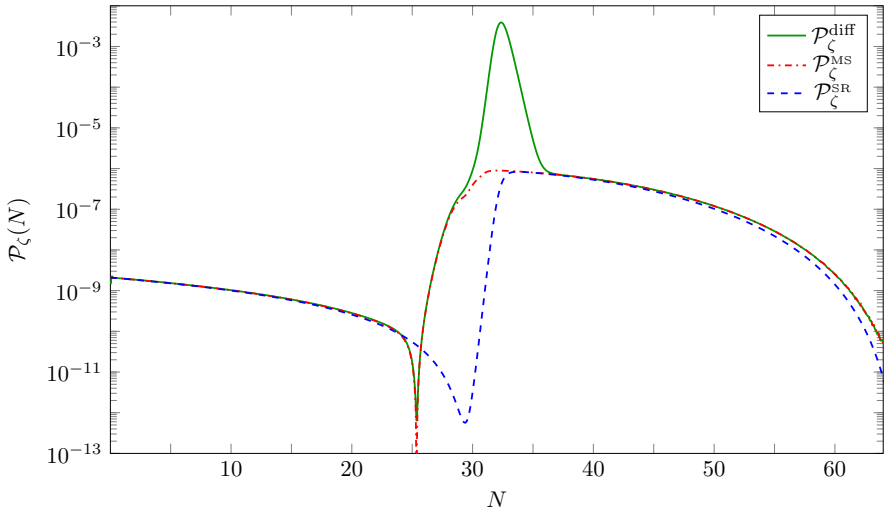

Figure 8.2. Power spectrum for the curvature perturbations $\mathcal{P}_\zeta$ computed from quantum diffusion (solid green), Eq. (8.2.12), solving the Mukhanov-Sasaki equation (red dash-dotted), Eq. (6.1.5), and using the slow-roll formula (blue dashed) in Eq. (6.1.6).

further modifications. In particular, we observe the peak resulting from the enhance of the noise beyond SR discussed before. This large enhancement of the fluctuations will certainly have relevant implications for the production of PBHs. Moreover, we observe that outside of the region of enhancement of the noise, $3-\epsilon+\epsilon_2<0$, the power spectrum computed with stochastic inflation recovers the numerical Mukhanov-Sasaki. This serves as a cross-check of the formalism.

8.2.2 Non-Gaussianities

One of the nice things of stochastic inflation is that it allows to compute higher order correlators of the coarse-grained field. The fact that there is a peak in the noise will also affect the non-Gaussian contributions to the PDF. We expect that any significant non-Gaussianity is localized around this region of large stochastic noise. We can solve any n-point function using the system of equations (8.2.7). In analogy to the second moments, the leading contribution to the 3-point functions is

$$\frac{d}{dN}\begin{pmatrix}\langle\delta\phi^3\rangle\\ \langle\delta\phi^2\delta\pi\rangle\\ \langle\delta\phi\delta\pi^2\rangle\\ \langle\delta\pi^3\rangle\end{pmatrix}=\begin{pmatrix}0 & -3 & 0 & 0\\ g(\epsilon_n) & -f(\epsilon_n) & -2 & 0\\ 0 & 2g(\epsilon_n) & -2f(\epsilon_n) & -1\\ 0 & 0 & 3g(\epsilon_n) & -3f(\epsilon_n)\end{pmatrix}\begin{pmatrix}\langle\delta\phi^3\rangle\\ \langle\delta\phi^2\delta\pi\rangle\\ \langle\delta\phi\delta\pi^2\rangle\\ \langle\delta\pi^3\rangle\end{pmatrix}+\begin{pmatrix}3\Xi_{\phi\phi}\langle\delta\phi\rangle\\ \Xi_{\phi\phi}\langle\delta\pi\rangle+\Xi^s_{\phi\pi}\langle\delta\phi\rangle\\ \Xi_{\pi\pi}\langle\delta\phi\rangle+\Xi^s_{\pi\phi}\langle\delta\pi\rangle\\ 3\Xi_{\pi\pi}\langle\delta\pi\rangle\end{pmatrix}. \tag{8.2.18}$$

A similar expression can be obtained for the 4-point function.

For the purpose of our analysis, we are interested in the Fourier transform of the

n-point functions. In particular, we are interested in the 3 and 4-point functions

$$\langle \zeta_k^3 \rangle = \frac{\kappa^3}{2^{3/2}} \frac{d^2}{d\ln k^2} \left(\frac{\langle \delta\phi^3 \rangle}{\epsilon^{3/2}} \right) \qquad \text{and} \qquad \langle \zeta_k^4 \rangle = \frac{\kappa^4}{4} \frac{d^3}{d\ln k^3} \left(\frac{\langle \delta\phi^4 \rangle}{\epsilon^2} \right) \tag{8.2.19}$$

which are associated respectively to the bispectrum and the trispectrum. Within this notation, $\langle \zeta_k^2 \rangle = \mathcal{P}_\zeta$. To measure the relative strength of these contributions one can normalize them with respect to the power spectrum, i.e. $\langle \bar{\zeta}_k^3 \rangle = \langle \zeta_k^3 \rangle / \langle \zeta_k^2 \rangle^{3/2}$ and $\langle \bar{\zeta}_k^4 \rangle = \langle \zeta_k^4 \rangle / \langle \zeta_k^2 \rangle^2$. In fact, these normalized moments will be the relevant quantities when computing the probability to form a PBH.

8.2.3 Characteristic function

In order to connect the n-point correlators with the PDF we can introduce the characteristic function $\chi(u_\phi, u_\pi)$, defined as the Fourier transform of the PDF

$$\chi(u_\phi, u_\pi) = \int_{-\infty}^{\infty} d\delta\phi \int_{-\infty}^{\infty} d\delta\pi e^{i(u_\phi \delta\phi + u_\pi \delta\pi)} P(N; \delta\phi, \delta\pi) \,, \tag{8.2.20}$$

where we are introducing now the PDF of the stochastic fluctuations $\delta\phi$ and $\delta\pi$. By denoting $\mathbf{u} = (u_\phi, u_\pi)$ and $\mathbf{x} = (\delta\phi, \delta\pi)$ we recover the generic multivariate definition of $\chi(\mathbf{u}) = \langle e^{i\mathbf{u}\cdot\mathbf{x}} \rangle$. The characteristic function is generated through the cumulant tensors $\kappa_{i_1 \cdots i_n}$ defined by

$$\kappa_{i_1 \cdots i_n} = (-i)^n \, \frac{\partial}{\partial u_1} \cdots \frac{\partial}{\partial u_n} \log \langle e^{i\mathbf{u}\cdot\mathbf{x}} \rangle \bigg|_{\mathbf{u}=0} \,. \tag{8.2.21}$$

In our case, the indices i_n could take only two values, either ϕ or π.

The cumulant tensors can be generically related to the n-point correlation functions. However, as it will be clarified in the next section, we will only be interested in the case in which we marginalize over the momentum. For the generic n-variate case, this is equivalent to marginalize over $n-1$ variables, which effectively leaves a univariate distribution with characteristic function $\chi(u_1) = \chi(\mathbf{u})|_{u_{i>1}=0}$. For our problem, we will have

$$\chi(u_\phi) = \langle e^{i u_\phi \delta\phi} \rangle = \exp\left[\sum_{p=1}^{\infty} \frac{i^p}{p!} \kappa_p u_\phi^p \right] \,, \tag{8.2.22}$$

where the first cumulants are given by

$$\begin{aligned}
\kappa_1 &= \langle \delta\phi \rangle \\
\kappa_2 &= \langle \delta\phi^2 \rangle - \langle \delta\phi \rangle^2 \\
\kappa_3 &= \langle \delta\phi^3 \rangle - 3\langle \delta\phi \rangle \langle \delta\phi^2 \rangle + 2\langle \delta\phi \rangle^3 \\
\kappa_4 &= \langle \delta\phi^4 \rangle - 4\langle \delta\phi \rangle \langle \delta\phi^3 \rangle + 12\langle \delta\phi \rangle^2 \langle \delta\phi^2 \rangle - 3\langle \delta\phi^2 \rangle^2 - 6\langle \delta\phi \rangle^4 \,.
\end{aligned} \tag{8.2.23}$$

With all these statistical machinery, we are now ready to compute the probability to form a PBH. Before doing so, let us briefly review the effect of the third and fourth moments in the PDF. For illustrative purposes, we have compared a Gaussian distribution ($\kappa_{n>2} = 0$)

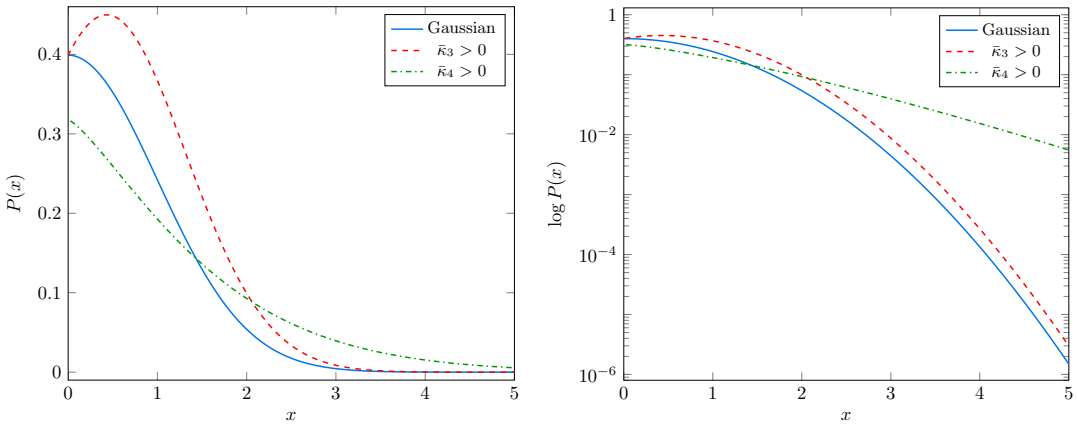

Figure 8.3. Illustration of the effect of a positive skewness, $\bar{\kappa}_3$, and a positive (excess) kurtosis, $\bar{\kappa}_4$, for the probability density function $P(x)$ (left) and its logarithm (right) in comparison with a Gaussian distribution ($\kappa_{n>2} = 0$). The dominant effect in the tail is given by a positive kurtosis.

to a PDF with $\kappa_3, \kappa_4 \neq 0$ in the left panel of Fig. 8.3 (see specific formula we are plotting in (C.2.5)). On the one hand, the third moment can induce an asymmetry of the PDF with respect to its mean value. This is commonly characterized by the skewness, which in terms of the cumulants reads $\bar{\kappa}_3 = \kappa_3/\kappa_2^{3/2}$. On the other hand, the fourth moment tell us how relevant are the tails with respect to the peak and it is measured by the (excess) kurtosis $\bar{\kappa}_4 = \kappa_4/\kappa_2^2$. For us, since we are interested in the probability above a certain threshold, the important part will be the tails of the PDF that can be better seen plotting the logarithm of the PDF as in the right panel of Fig. 8.3. There, one can clearly see that both a positive skewness and kurtosis induce higher tails compared to the Gaussian. For our toy model, we find that the tails of the PDF are dominated by a positive kurtosis. This will enlarge the probability to form a PBH. However, one should remember that non-Gaussianities could also reduce the fraction of PBH if the distribution is characterized by a negative skewness or kurtosis.

8.3 Implications for PBH production

PBH are formed when large fluctuations produced during inflation re-enter the horizon and collapse all the content within a causal horizon. The fraction of PBH formed thus depends on the probability to have fluctuations larger than a certain threshold at the scale of the horizon. The probability distribution in single-field inflation is a bivariate distribution that depends on the fluctuations and their velocities. Therefore, the fraction of PBH formed is computed from

$$\beta_f(M) = P(\zeta > \zeta_c) = \int_{\zeta_c}^{\infty} \int_{-\infty}^{\infty} P(M; \zeta, \delta\pi) d\delta\pi d\zeta \,, \tag{8.3.1}$$

where we have integrated over the momentum fluctuations.[5]

Altogether, the abundance of PBHs will be sensitive to the PDF of the stochastic fluctuations. We will show next how this abundance can be connected with the correlation functions of these fluctuations. For the case of Gaussian fluctuations then all the information is contained in the 2-point functions that can be related to the power spectrum P_ζ. A peak in P_ζ thus produce a higher probability to form PBHs. This is the whole purpose of considering inflation with a quasi-inflection point as a source of PBHs. However, if the fluctuations are non-Gaussian (as it will be in our case), these additional contributions from higher n-point functions will be relevant. This is because we are interested in the probability of being above a certain threshold and, as a consequence, we are very sensitive to the tails of the distribution. The fact that non-Gaussianities can be relevant in the production of PBHs has already been considered in [424,465–470]. When considering quantum diffusion in SR inflation, non-Gaussianities played a central role [471–474]. Here, for quantum diffusion beyond SR, non-Gaussianities will as well be very important.

8.3.1 Abundance of PBH

The abundance of PBH is then determined by $P(\zeta > \zeta_c) = \langle \Theta(\zeta - \zeta_c) \rangle_P$. We can express this probability in terms of the characteristic function by

$$P(\zeta > \zeta_c) = \int_{-\infty}^{\infty} d\zeta \int_{\zeta_c}^{\infty} da \int_{-\infty}^{\infty} \frac{du}{2\pi} e^{iu(\zeta - a)} P(\zeta) = \int_{\zeta_c}^{\infty} da \int_{-\infty}^{\infty} \frac{du}{2\pi} e^{-iua} \chi(u) \,, \quad (8.3.2)$$

where in the first equality we have introduce the integral definition of the step function[6]. Therefore, as anticipated in Sec. 8.2.3, this probability is only sensitive to the marginalized characteristic function $\chi(u)$. Such function can be constructed with the cumulants of the PDF, cf. (8.2.22), which can be obtained solving the evolution of the correlations functions (8.2.7) as explained before.

Conveniently, the integral of $P(\zeta > \zeta_c)$ can be solved exactly. The resulting formula is

$$P(\zeta > \zeta_c) = \frac{1}{2} \mathrm{Erfc}\left[\frac{z_c}{\sqrt{2}}\right] + \frac{e^{-z_c^2/2}}{\sqrt{\pi}} \sum_{n=3}^{\infty} \frac{\bar{K}_n}{2^{n/2} n!} H_{n-1}\left[\frac{z_c}{\sqrt{2}}\right] \,, \quad (8.3.3)$$

where Erfc is the complementary error function and H_n are the Hermite polynomial. The normalized threshold is given by $z_c = (\zeta_c - \kappa_1)/\kappa_2^{1/2}$, where we recall that κ_n are the cumulants of the PDF. The functions $\bar{K}_n$ can be related to the normalized cumulants

[5] One may wonder if the gravitational collapse could depend on the canonical momentum fluctuations $\delta\pi$, which could happen if entropy as well as adiabatic density fluctuations existed at horizon re-entry. However, although entropy fluctuations can be excited during the non-slow-roll phase, only adiabatic perturbations are present at the end of inflation. This means that only curvature fluctuations will re-enter the horizon and, as a consequence, whether PBH are formed or not depends only on the size of curvature gradients. That is the reason why we must marginalize over the canonical momentum fluctuations. Another issue might be the effect of non-Gaussianities in the formation of PBH, which has not been studied yet in depth in the literature [405], and would require a dedicated study.

[6] The step function can be written as $\Theta(x - x_c) = \int_{x_c}^{\infty} da \int_{-\infty}^{\infty} \frac{du}{2\pi} e^{iu(x-a)}$.

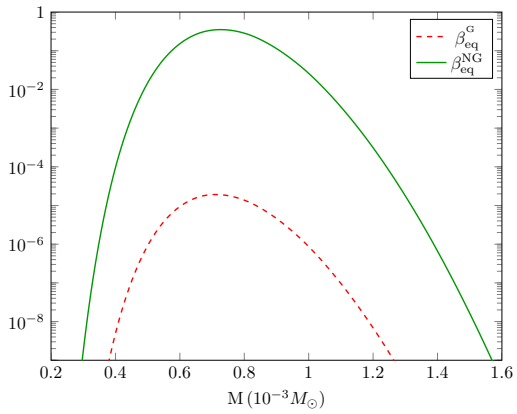

Figure 8.4. Fraction of PBH at the time of equality $\beta_{\rm eq}$ as a function their mass computed using quantum diffusion and fixing $\zeta_c = 0.5$. The abundance given by the Gaussian approximation is compared with the non-Gaussian, which is orders of magnitude larger.

$\bar{\kappa}_n$ and only differ from them for $n \geqslant 6$. This result was originally obtained in Ref. [464] using the path integral formalism and recently revisited in [470]. Since it is an important formula, we summarize for completeness the derivation of this result and the details in the definitions in App. C.4 using the language of the characteristic function and cumulants that we are using here.

Regarding the effects of quantum diffusion on the abundance of PBH, there will be two effects. First, the fact that the slow-roll violation enhances quantum fluctuations that back-react producing an additional peak (on top of the classical one) in the power spectrum will significantly increase the number of PBH that are formed. This can be very easily understood looking at the first term of (8.3.3), which accounts for the Gaussian contribution. The threshold $z_c^{\rm diff}$ for quantum diffusion will be lowered with respect to the usual one $z_c^{\rm MS}$ by

$$z_c^{\rm diff} \simeq \left(\frac{\mathcal{P}_\zeta^{\rm MS}}{\mathcal{P}_\zeta^{\rm diff}}\right)^{1/2} z_c^{\rm MS} \,. \tag{8.3.4}$$

Depending on the size of the diffusion peak, the difference around this point could be of several orders of magnitude. In Fig. 8.4, we present the fraction of PBH at equality $\beta_{\rm eq}^{\rm G}$ for a Gaussian distribution with variance given by the power spectrum computed using stochastic inflation $(\mathcal{P}_\zeta^{\rm diff})^{1/2}$ (cf. Fig. 8.2) and $\zeta_c = 0.5$. The corresponding fraction of PBH without considering the diffusion, computed from $(\mathcal{P}_\zeta^{\rm MS})^{1/2}$, is very suppressed, due to several orders of magnitude difference in the peak, and does not appear in the plot.

Second, non-Gaussiniaties will change the above prediction by either augmenting or decreasing the number of PBH depending on the sign of each contribution. For our case of study, we solve the stochastic evolution up to fourth order.[7] We find that the

[7]We have checked that the fifth order contribution is negligible. However, we realize that in order to derive the sixth order reliably we require higher time resolution than the one used in this analysis. An

fourth moment is the dominant piece. It contributes with a positive large kurtosis. This significantly amplifies the production of PBHs as it can be seen from the fraction of PBH at equality $\beta_{\rm eq}^{\rm NG}$ plotted in Fig. 8.4. In this example, we have chosen the parameters so that the energy density of PBH at equality $\Omega_{\rm PBH}^{\rm eq}$, when considering the NG contributions, is the dominant fraction of the DM. Considering only the Gaussian contribution would lead to PBHs being a very subdominant fraction of DM. This shows that indeed NG are very relevant and they could soften the necessity to have very large peaks in the power spectrum. In the following, we discuss the implications of these results for inflationary model building and how they generalize.

8.3.2 Model building

In the previous section, we have found that in models of inflation with a quasi-inflection point quantum diffusion causes two main effects: *i*) there is an enhancement of the power spectrum and *ii*) non-Gaussian corrections become relevant, modifying the prediction for the abundance of PBHs. These results have profound implications when constructing models to produce PBHs. On the one side, there will be regions of the parameter space of previously considered models in the literature now excluded because they would overproduce PBHs. On the other side, there will be other regions, possibly with less tuning of the parameters, that now would copiously produce PBHs.

For the purpose of this work, we have not performed a detailed analysis of the viable parameter space for inflationary models with a near-inflection point accounting for quantum diffusion. Such re-analysis should be performed elsewhere. Our expectation is that any quantum diffusion effect would be closely related to the type of deviation from slow-roll. For example, we find that the height of the diffusion peak is linked to how much the inflaton slows-down, i.e. how small ϵ becomes, while its width depends on how fast the transition from slow-roll to ultra slow-roll happens, i.e. for how long $\epsilon_2 < 0$. In any case, we observe that the violation of SR should be such that during an interval $3 - \epsilon + \epsilon_2 < 0$ so that there is an amplification of the stochastic noise. Moreover, we encounter that the enhancement of the power spectrum and the relevance of non-Gaussian corrections are highly correlated. This implies that the larger the diffusion peak is, the more the Gaussian prediction for the abundance of PBHs will be modified. For the toy model potential that we have considered, we find that a positive kurtosis dominates over the skewness to enlarge the production of PBHs. From these considerations, we expect that quantum diffusion would have an important role in models such a critical Higgs inflation [4], single-field double inflation [412], radiative plateau inflation [415] or some string inspired realizations [416–418].

alternative would be to compute directly the tail of the PDF expanding around large ζ. We will develop a method to do so in chapter 9.

CHAPTER

9

The tail of curvature perturbation distributions

The abundance of primordial black-holes is determined by the probability that a curvature perturbation ζ exceeds a given critical value ζ_c (recall (6.0.1)). Since we are integrating above a certain threshold $\zeta_c \sim 1$, the fraction of PBHs will be dominated by the tail of PDF, being exponentially sensitive to its decay rate. We have seen in the previous chapter 8 that quantum diffusion generically introduces non-Gaussian corrections. Corrections that are most relevant in the tail of the PDF. In this chapter we develop novel techniques to compute directly the tail of the PDF to properly assess the abundance of PBHs.

Let $P_\phi(\mathcal{N})$ be the probability that, starting from the initial field value ϕ, $\mathcal{N}$ inflationary e-folds are realised. Here, ϕ is a generic field space coordinate, i.e. it has several components in the case of multiple-field inflation or if one does not assume slow roll and the dynamics of velocity coordinates should also be incorporated. In Ref. [474], it is shown that $P_\phi(\mathcal{N})$ obeys the partial differential equation

$$\frac{\partial P_\phi(\mathcal{N})}{\partial \mathcal{N}} = \mathcal{L}^\dagger_{\rm FP}\left(\phi\right) \cdot P_\phi(\mathcal{N}), \tag{9.0.1}$$

where $\mathcal{L}^\dagger_{\rm FP}$ is the adjoint of the Fokker-Planck operator that drives the stochastic evolution of the system. In other words, if one denotes by $P_N\left(\phi\right)$ the probability that, after N e-folds, the system lies in the configuration ϕ (staring from some initial conditions), this probability is driven by the Fokker-Planck equation

$$\frac{\partial P_N\left(\phi\right)}{\partial N} = \mathcal{L}_{\rm FP}\left(\phi\right) \cdot P_N\left(\phi\right) \ . \tag{9.0.2}$$

One should note that although $P_\phi(\mathcal{N})$ and $P_N\left(\phi\right)$ are conceptually very different objects, they obey similar equations. We are going to use this fact in what follows.

In contrast to chapter 8, as a first step, we will limit our analysis to slow-roll inflation. In such case, for a inflationary potential $V(\phi)$, and denoting

$$v = \frac{V}{24\pi^2 M_{\rm pl}^4} \tag{9.0.3}$$

the reduced potential, the Fokker-Planck operator and its adjoint are given by [480]

$$\frac{1}{M_{\rm pl}^2}\mathcal{L}_{\rm FP} \quad = \sum_i \frac{\partial}{\partial\phi_i}\frac{v_{\phi_i}}{v} + \sum_i \frac{\partial^2}{\partial\phi_i^2} v\,, \tag{9.0.4}$$

$$\frac{1}{M_{\rm pl}^2}\mathcal{L}^\dagger_{\rm FP} \quad = -\sum_i \frac{v_{\phi_i}}{v}\frac{\partial}{\partial\phi_i} + v\sum_i \frac{\partial^2}{\partial\phi_i^2}\,. \tag{9.0.5}$$

$$\tag{9.0.6}$$

In order to obtain the behavior of the tail of the PDF, we can make an expansion around $\mathcal{N} = \infty$ via

$$P_\phi(\mathcal{N}) = \sum_n a_n(\phi) e^{-\Lambda_n \mathcal{N}}\,, \tag{9.0.7}$$

where the coefficients $a_n(\phi)$ determine the amplitude of the tail and Λ_n the exponential decay. Notice that this expansion is different from the classical picture where one typically assumes a Gaussian PDF

$$P_\phi(\mathcal{N})|_{\rm cl} \propto \exp\left(-\frac{1}{2}\frac{\mathcal{N}^2}{\mathcal{P}_{\zeta,\rm cl}}\right)\,, \tag{9.0.8}$$

where $\mathcal{P}_{\zeta,\rm cl} = 2v^3/(M_{\rm pl}^2 v'^2)$ is the classical value of the power spectrum. Therefore, although at small values of $\mathcal{N}$ the classical and stochastic PDFs might give similar results, their tails differ substantially. We are now going two present two complementary techniques to compute $a_n(\phi)$ and Λ_n.

9.1 Finding the poles of the characteristic function

A fundamental measure of the statistical properties of the curvature perturbations is the characteristic function, introduced in section 8.2.3 and defined through

$$\chi_\mathcal{N}(t,\phi) \equiv \left\langle e^{it\mathcal{N}(\phi)}\right\rangle\,, \tag{9.1.1}$$

where t is an auxiliary parameter. This function is nothing but the Fourier transform of the PDF,

$$P_\phi(\mathcal{N}) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-it\mathcal{N}}\chi_\mathcal{N}(t,\phi)dt\,, \tag{9.1.2}$$

and satisfies the differential equation

$$\left(\mathcal{L}^\dagger_{\rm FP} + it\right)\chi_\mathcal{N}(t,\phi) = 0 \tag{9.1.3}$$

with boundary conditions

$$\chi_\mathcal{N}(t,\phi_{\rm end}) = 1 \quad \text{and} \quad \chi'_\mathcal{N}(t,\phi_{\rm uv}) = 0\,, \tag{9.1.4}$$

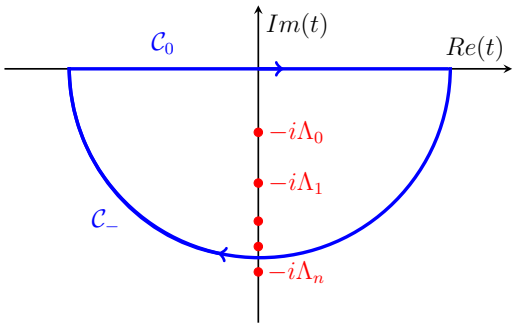

Figure 9.1. Schematic representation of the pole structure of the characteristic function. We can extend the real axis integral of $P_\phi(\mathcal{N})$ in (9.1.7) using the contour in the complex plane $\mathcal{C}_0 \cup \mathcal{C}_-$.

where $\phi_{\rm end}$ is the value of ϕ at the end of inflation and $\phi_{\rm uv}$ the largest initial value.

We can use the analytic structure of the characteristic function to obtain the tail of the PDF. In particular, from the integral above, we notice that only $\chi_{\mathcal{N}}(t,\phi)$ could be singular. Therefore, we could solve the integral over the real axis extending it to the complex plane and using a contour as in Fig. 9.1. Then, applying the residue theorem,

$$\oint \frac{f(z)\,dz}{(z-z_p)^{n+1}} = \frac{2\pi i}{n!}\left[\frac{d^n}{dz^n} f(z)\right]_{z=z_p}\,, \tag{9.1.5}$$

we find that the tail expansion (9.0.7) is controlled by the poles of the characteristic function on the negative imaginary axis $t = -i\Lambda_n$. This can be easily seen if we isolate the singularities

$$\chi_{\mathcal{N}}(t,\phi) = \sum_n \frac{a_n(\phi)}{\Lambda_n - it} + g(t,\phi)\,, \tag{9.1.6}$$

where $g(t,\phi)$ is a regular function. Note that we are considering only simple poles, assumption that we have verified in practice numerically. Moreover, since the Fokker-Planck operator and its adjoint are positive operators, so the Λ_n are all positive. Accordingly, integrating the PDF over the contour $\mathcal{C} = \mathcal{C}_0 \cup \mathcal{C}_-$ in Fig. 9.1 we obtain

$$\begin{aligned} P_\phi(\mathcal{N}) &= \frac{1}{2\pi}\oint_{\mathcal{C}} \sum_n \frac{a_n(\phi)\, e^{-it\mathcal{N}}}{\Lambda_n - it} dt + \frac{1}{2\pi}\oint_{\mathcal{C}} g(t,\phi)\, e^{-it\mathcal{N}} dt \\ &= \sum_n a_n(\phi)\, e^{-\Lambda_n \mathcal{N}}\,. \end{aligned} \tag{9.1.7}$$

Therefore, the exponential decay of the PDF at large $\mathcal{N}$ is governed by the lowest pole of the characteristic function Λ_0 and its residue $a_0(\phi)$. In practice, the lowest pole can be obtained by solving the characteristic function equation (9.1.3) and finding the zeros of its inverse $\chi^{-1}_{\mathcal{N}}(t,\phi)$. The residue is just the derivative of the characteristic function w.r.t. t at that point $t = -i\Lambda_0$. An important remark is that the value of the pole is independent of the initial field value ϕ while the residue does depend. This implies that the decay rate of the PDF tail is universal for a given potential.

9.2 An equivalent eigenvalue problem

Now, since the equation (9.0.1) that governs $P_\phi(\mathcal{N})$ is similar to the diffusion equation (9.0.2), one can employ late-time limit techniques. At the formal level, this works as follows. One first notices that Eq. (9.0.1) can be formally solved as

$$P_\phi\left(\mathcal{N}\right) = \exp\left[\mathcal{N}\mathcal{L}^\dagger_{\rm FP}\left(\phi\right)\right] P_\phi\left(\mathcal{N}=0\right) \,. \tag{9.2.1}$$

One then introduces the orthonormal set of eigenfunctions Φ_n of the operator $\mathcal{L}^\dagger_{\rm FP}$ that satisfy the boundary conditions of the problem,

$$\mathcal{L}^\dagger_{\rm FP}\cdot\Phi_n\left(\phi\right) = -\Lambda_n\Phi_n\left(\phi\right) \tag{9.2.2}$$

(here a minus sign is introduced in the eigenvalue for notational convenience) and decomposes $P_\phi\left(\mathcal{N}=0\right)$ on the basis formed by these functions,

$$P_\phi\left(\mathcal{N}=0\right) = \sum_n \alpha_n\Phi_n\left(\phi\right) \,. \tag{9.2.3}$$

This gives rise to

$$P_\phi\left(\mathcal{N}\right) = \sum_n \alpha_n\Phi_n\left(\phi\right) e^{-\Lambda_n\mathcal{N}} \tag{9.2.4}$$

If one orders the eigenvalues, $0 \leqslant \Lambda_0 < \Lambda_1 < \Lambda_2 < \cdots < \Lambda_n$, the above expression is nothing but a tail expansion of the distribution $P_\phi\left(\mathcal{N}\right)$, equivalent to (9.0.7) with $a_n(\phi) = \alpha_n\Phi_n(\phi)$.

To get the dominant behavior, one therefore simply needs to solve the eigenvalue problem

$$\Phi_n'' - \frac{v_\phi}{v^2}\Phi_n' + \frac{\Lambda_n}{vM_{\rm pl}}\Phi_n = 0 \tag{9.2.5}$$

with boundary conditions

$$\Phi_n(\phi_{\rm end}) = \Phi_n'(\phi_{\rm uv}) = 0\,. \tag{9.2.6}$$

The first boundary condition comes from the fact that $P_{\phi_{\rm end}}(\mathcal{N}) = \delta\left(\mathcal{N}\right)$, so all eigen components should be identically zero except when $\Lambda_n = \infty$. The second boundary condition comes from the reflective wall located at $\phi_{\rm uv}$. One could notice that the above equation is the same as the one for the characteristic function (9.1.3) changing $t \to -i\Lambda_n$. However, the boundary conditions are different.

The coefficients in the expansion (9.0.7), α_n, can be determined as follows. One first notes that the functions Φ_n form an orthogonal set, in the sense that

$$\langle\Phi_n, \Phi_m\rangle \equiv \int_{\phi_{\rm end}}^{\phi_{\rm uv}} \Phi_n\left(\phi\right)\Phi_m\left(\phi\right){\rm d}\phi = \frac{\Delta\phi}{2}\delta_{n,m}\,, \tag{9.2.7}$$

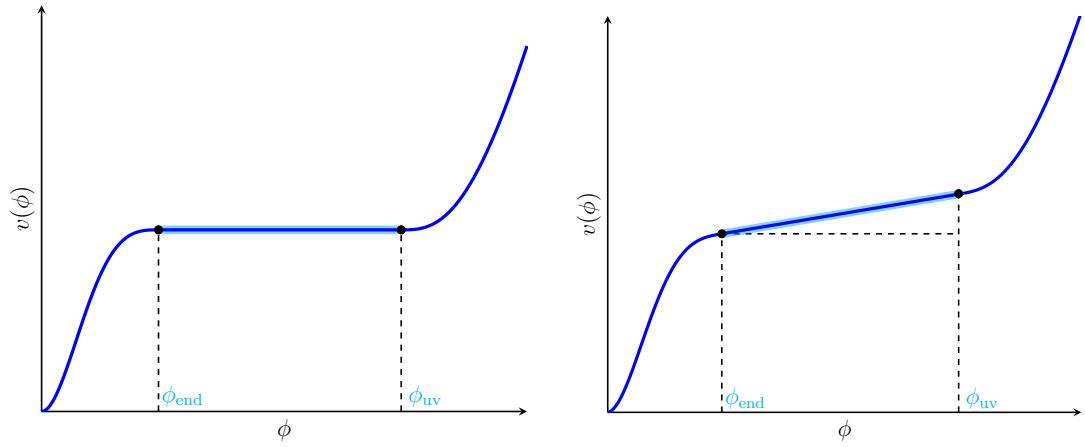

Figure 9.2. Schematic representation of the flat (left) and linear (right) potentials studied in sections 9.3 and 9.3 respectively. We only consider the region between $\phi_{\rm end}$ and $\phi_{\rm uv}$.

where $\Delta\phi = \phi_{\rm end} - \phi_{\rm uv}$. Making use of the expansion (9.0.7), this gives rise to

$$\langle \Phi_n, P_\phi(\mathcal{N}) \rangle = \alpha_n \frac{\Delta\phi}{2} e^{-\Lambda_n \mathcal{N}} \,. \tag{9.2.8}$$

We can compare this expression with the one obtained from the characteristic function,

$$\begin{aligned} \langle \Phi_n, P_\phi(\mathcal{N}) \rangle &= \tfrac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_{\phi_{\rm end}}^{\phi_{\rm uv}} \Phi_n(\phi)\, \chi_{\mathcal{N}}(t,\phi)\, \mathrm{d}\phi \right] e^{-it\mathcal{N}} \mathrm{d}t & (9.2.9)\\ &= \tfrac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_{\phi_{\rm end}}^{\phi_{\rm uv}} \Phi_n(\phi)\, a_n(\phi) \mathrm{d}\phi \right] \tfrac{e^{-it\mathcal{N}}}{\Lambda_n - it} \mathrm{d}t & (9.2.10)\\ &= \left[\int_{\phi_{\rm end}}^{\phi_{\rm uv}} \Phi_n(\phi)\, a_n(\phi) \mathrm{d}\phi \right] e^{-\Lambda_n \mathcal{N}} & (9.2.11) \end{aligned}$$

where we have used first the orthogonality of $\Phi_n(\phi)$ and then the residue theorem. By identifying Eqs. (9.3.27) and (9.3.10), one obtains

$$\alpha_n = \frac{2}{\Delta\phi} \left[\int_{\phi_{\rm end}}^{\phi_{\rm uv}} \Phi_n(\phi)\, a_n(\phi) \mathrm{d}\phi \right] \,. \tag{9.2.12}$$

We will use this expression in next section.

9.3 Applications for inflationary potentials

In the following we are going to apply the two techniques developed in the previous sections to compute the tail of the PDF in different types of potentials.

Quantum well potential

Let us begin by considering a potential that is exactly flat, $v = v_0$, between $\phi_{\rm end} = 0$ and $\phi_{\rm uv} = \Delta\phi_{\rm well}$ as in the left panel of Fig. 9.2. As we will see later, this type of potentials is interesting because it serves as a good effective description of potentials of the form $v = v_0[1 + (\phi/\phi_0)^p]$, for which $\Delta\phi_{\rm well} = \phi_0 v_0^{1/p}$, and where the classical part of the potential above $\Delta\phi_{\rm well}$ acts as a reflective wall.

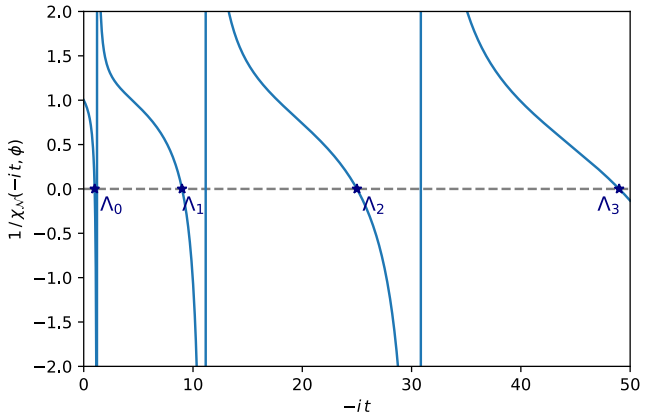

Figure 9.3. Zeros of the inverse characteristic function for a flat potential. We have chosen $\mu^2 = \pi^2/4$ so that zeros are located at $\Lambda_n = (2n+1)^2$. We have evaluated $\chi_{\mathcal{N}}^{-1}(t,\phi)$ at $\phi = \Delta\phi_{\rm well}/10$.

Poles of the characteristic function. In this simple example, the equation for the characteristic function (9.1.3) is just

$$\chi''_{\mathcal{N}}(t,\phi) + \frac{i\,t}{v_0 M_{\rm pl}^2}\chi_{\mathcal{N}}(t,\phi) = 0\,. \tag{9.3.1}$$

This equation has an analytic solution

$$\chi_{\mathcal{N}}(t,\phi) = \frac{\cos\left[(it)^{1/2}\mu(x-x_{\rm uv})\right]}{\cos\left[(it)^{1/2}\mu\right]}\,, \tag{9.3.2}$$

where we have defined $x = \phi/\Delta\phi$ and introduced the quantity

$$\mu^2 = \frac{\Delta\phi^2}{v_0 M_{\rm pl}^2} \tag{9.3.3}$$

as in Ref. [474]. When $\phi_{\rm end} = 0$, then $\Delta\phi = \phi_{\rm uv} \equiv \Delta\phi_{\rm well}$ and the expression above simplifies with $x_{\rm uv} = 1$. Replacing $it \to \Lambda_n$, we find that $\chi_{\mathcal{N}}(it,\phi)$ is a real function and has periodic poles at

$$\Lambda_n = \frac{\pi^2}{\mu^2}\left(n+\frac{1}{2}\right)^2\,. \tag{9.3.4}$$

Therefore, the exponential decay of the tail of the PDF depends both on the size of the quantum well $\Delta\phi_{\rm well}$ and its scale v_0. We have plotted the inverse characteristic function for a flat potential in Fig. 9.3. The zeros of $\chi_{\mathcal{N}}^{-1}(t,\phi)$ correspond to $-i\Lambda_n$.

Eigenvalue problem. In this case the eigenvalue problem (9.2.5) reads

$$\Phi''_n(\phi) + \frac{\Lambda_n}{v_0 M_{\rm pl}^2}\Phi_n(\phi) = 0, \tag{9.3.5}$$

with boundary conditions $\Phi_n(0) = \Phi_n'(\Delta\phi_{\rm well}) = 0$. The generic solution of the above equation is

$$\Phi_n\left(\phi\right) = A_n \exp\left(i\sqrt{\frac{\Lambda_n}{v_0 M_{\rm pl}^2}}\phi\right) + B_n \exp\left(-i\sqrt{\frac{\Lambda_n}{v_0 M_{\rm pl}^2}}\phi\right) . \tag{9.3.6}$$

The first boundary condition $\Phi_n(0) = 0$ imposes that $B_n = -A_n$, hence

$$\Phi_n\left(\phi\right) \propto \sin\left(\sqrt{\frac{\Lambda_n}{v_0 M_{\rm pl}^2}}\phi\right) . \tag{9.3.7}$$

The second boundary condition $\Phi_n'(\Delta\phi_{\rm well}) = 0$ then implies that

$$\cos\left(\sqrt{\frac{\Lambda_n}{v_0 M_{\rm pl}^2}}\Delta\phi_{\rm well}\right) = 0\,, \tag{9.3.8}$$

i.e. one obtains again (9.3.4). One can check that the above expression implies that all Λ_n's are positive, in agreement with the fact that the adjoint Fokker-Planck operator is indeed positive. One then has

$$\Phi_n\left(\phi\right) = \sin\left[\pi\left(n+\frac{1}{2}\right)\frac{\phi}{\phi_{\rm UV}}\right] \tag{9.3.9}$$

and the expansion (9.0.7) is in perfect agreement with Eq. (4.10) of Ref. [474]. In particular, the dominant behaviour $\propto e^{-\pi^2\mathcal{N}/(4\mu^2)}$ is correctly reproduced, but all higher orders too. We can compute the coefficients α_n using the solution for the characteristic function and the procedure outlined previously, cf. (9.2.12). One obtains

$$\langle\Phi_n, P_\phi\left(\mathcal{N}\right)\rangle \quad = \tfrac{\pi\phi_{\rm UV}}{\mu^2}\left(n+\tfrac{1}{2}\right)e^{-\Lambda_n\mathcal{N}}\,. \tag{9.3.10}$$

By identifying Eqs. (9.3.27) and (9.3.10), one finds

$$\alpha_n = \frac{2\pi}{\mu^2}\left(n+\frac{1}{2}\right) , \tag{9.3.11}$$

which exactly matches Eq. (4.10) of Ref. [474], but which here we have derived in another, slightly more straightforward, way. Altogether, the PDF for the constant potential can be written as

$$P_\phi(\mathcal{N}) = \sum_n \frac{2\pi}{\mu^2}\left(n+\frac{1}{2}\right)\sin\left(\sqrt{\frac{\Lambda_n}{v_0 M_{\rm pl}^2}}\phi\right)\,e^{-\Lambda_n\mathcal{N}}\,. \tag{9.3.12}$$

In Fig. 9.4 we plot both the full PDF and the leading term in the tail expansion $a_\phi(\phi)e^{-\Lambda_n\mathcal{N}}$. As it can be observed, at large $\mathcal{N}$ the tail expansion is a very good approximation. Note also that we have rescaled the number of e-folds by $1/\mu^2$ as well as the PDF. Small values of v_0 thus imply large $\mathcal{N}$ and $P_\phi(\mathcal{N})$.

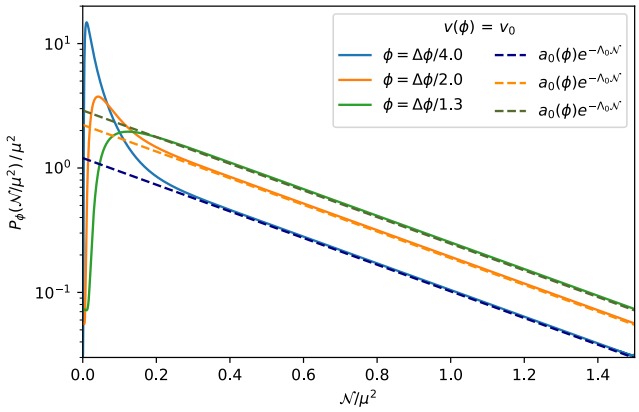

Figure 9.4. Probability distribution function of the number of e-folds $\mathcal{N}$ starting at different initial field values ϕ for a flat potential. We compare the PDF (solid) with the leading term of the tail expansion (dashed), rescaling $\mathcal{N}$ by $\mu^2 = \Delta\phi^2/v_0$.

Vacuum dominated potential with constant slope

Let us now consider a potential of the type

$$v\left(\phi\right) = v_0 \left(1 + \alpha \frac{\phi}{M_{\rm pl}}\right) \tag{9.3.13}$$

with constant slope α. Since $\epsilon_1 \simeq \alpha^2/2$, one should have $\alpha \ll 1$ in order for slow roll to be valid. Moreover, we will consider only scenarios between $\phi = 0$ and $\phi_{\rm UV} \ll M_{\rm pl}/\alpha$ where the potential is almost constant $v \simeq v_0$ (see right panel of Fig. 9.2).

Poles of the characteristic function. In this case the equation for the characteristic function (9.1.3) becomes

$$\chi''_{\mathcal{N}}(t,\phi) - \frac{v_0 \alpha}{M_{\rm pl} v(\phi)^2} \chi'_{\mathcal{N}}(t,\phi) + \frac{i\,t}{v(\phi)} \chi_{\mathcal{N}}(t,\phi) = 0\,, \tag{9.3.14}$$

where there is a friction term not present in the previous flat potential. More importantly, since the coefficients of the differential equation depend on ϕ, there is no generic analytic solution. There is, however, an analytic solution in the almost constant approximation in which one solves instead

$$\chi''_{\mathcal{N}}(t,\phi) - \frac{\alpha}{M_{\rm pl} v_0} \chi'_{\mathcal{N}}(t,\phi) + \frac{i\,t}{v_0^2} \chi_{\mathcal{N}}(t,\phi) = 0\,. \tag{9.3.15}$$

The solution reads

$$\chi_{\mathcal{N}}\left(t,\phi\right) = e^{\frac{\alpha\mu(x-x_{\rm end})}{2\sqrt{v_0}}} \frac{2\gamma\sqrt{it v_0}\cos\left[\mu\gamma\sqrt{it}(x-x_{\rm uv})\right] - \alpha\sin\left[\mu\gamma\sqrt{it}(x-x_{\rm uv})\right]}{2\gamma\sqrt{t v_0}\cos\left[\mu\gamma\sqrt{it}\right] + \alpha\sin\left[\mu\gamma\sqrt{it}\right]} \tag{9.3.16}$$

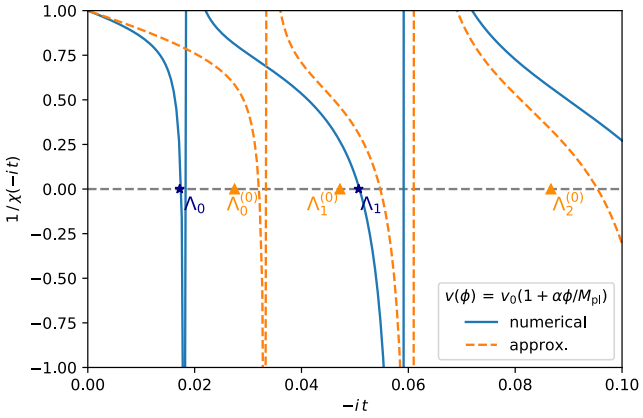

Figure 9.5. Zeros of the inverse characteristic function for a linear potential. We compare the numerical result of $\chi_{\mathcal{N}}^{-1}(t,\phi)$ (solid) and its poles Λ_n with the almost constant approximation (dashed) and $\Lambda_n^{(0)}$ given in (9.3.16) and (9.3.19) respectively. For better visualization, we have chosen $v_0 = 0.1$, $\alpha = 0.1$ and evaluated $\chi_{\mathcal{N}}^{-1}(t,\phi)$ at $\phi_{\rm uv}/10$. A smaller v_0 would give better agreement.

with

$$\gamma = \sqrt{1 - \frac{\alpha^2}{4itv_0}} \tag{9.3.17}$$

and $x = \phi/\Delta\phi$. When $\alpha = 0$, this boils down to the flat potential (9.3.2).

The poles of the characteristic function at $t = -i\Lambda_n$ are determined by the trascendental equation

$$\tan\left[\sqrt{\Lambda_n - \frac{\alpha^2}{4v_0}}\mu\right] = -\sqrt{\frac{4v_0}{\alpha^2}\Lambda_n - 1} \tag{9.3.18}$$

When v_0 is very small (in a sense that we will make more precise below), an approximate solution to the above equation is given by

$$\Lambda_n \simeq \Lambda_n^{(0)} \equiv \frac{\alpha^2}{4v_0} + \frac{\pi^2}{\mu^2}\left(n + \frac{1}{2}\right)^2 \tag{9.3.19}$$

where we have used the definition of μ in (9.3.3). Otherwise one has to solve a transcendental equation. One can see that the condition for this approximation to hold is that

$$v_0 \ll \alpha\frac{\phi_{\rm uv}}{M_{\rm pl}}\,. \tag{9.3.20}$$

If one takes $\phi_{\rm uv}$ to its maximal allowed value, $\phi_{\rm uv} \sim M_{\rm pl}/\alpha$, this is satisfied as soon as $v_0 \ll 1$, which is always the case. In Fig. 9.5 we show the inverse characteristic function

obtained by solving numerically (9.3.14) (solid line) and the analytical solution (9.3.16) in the almost constant approximation (dashed line). We also include the approximate values of $\Lambda_n^{(0)}$ given in (9.3.19). To improve the visualization of the different poles and approximations, we choose $v_0 = 0.1$ and $\alpha = 0.1$. Having a smaller v_0 increases the agreement between the three cases but also reduces the separation between Λ_n. As in the constant potential case, changing the value of $\phi_{\rm uv}$ affects the location of the poles.

In the small v_0 limit, one can see that the effect of adding a slope in the potential is simply to shift all eigenvalues by a fixed quantity, namely $\alpha^2/(4v_0)$. However, this shift can be very important for the first eigenvalues. Indeed, if one pushes $\phi_{\rm uv}$ to the largest values it can take for the vacum dominated approximation to be valid, namely $\phi_{\rm uv} \sim M_{\rm pl}/\alpha$, then $1/\mu^2$ is of order $\alpha^2 v_0$, which is much smaller than α^2/v_0 if one works at sub-Planckian energies $v_0 \ll M_{\rm pl}$.

In this limit, the dominant behavior on the tail is therefore given by $P_\phi(\mathcal{N}) \simeq \exp\left[-\frac{\alpha^2}{4v_0}\mathcal{N}\right]$, and the exponential suppression is much stronger than in the pure diffusive limit. However, there are many eigenvalues clustered close to this dominant branch and this quasi continuum might lead to an enhancement of the resulting power. Furthermore, let us notice that the classical value of the power spectrum, $\mathcal{P}_{\zeta,\rm cl} = 2v^3/(M_{\rm pl}^2 v'^2)$, in this model is given by $\mathcal{P}_{\zeta,\rm cl} \simeq 2v_0/\alpha^2$. This means that the dominant behavior on the tail can be written as

$$P_\phi(\mathcal{N}) \underset{\mathcal{N}\gg 1}{\propto} \exp\left(-\frac{1}{2}\frac{\mathcal{N}}{\mathcal{P}_{\zeta,\rm cl}}\right) \tag{9.3.21}$$

while in the classical picture, the distribution is Gaussian and one typically assumes

$$P_\phi(\mathcal{N})|_{\rm cl} \underset{\mathcal{N}\gg 1}{\propto} \exp\left(-\frac{1}{2}\frac{\mathcal{N}^2}{\mathcal{P}_{\zeta,\rm cl}}\right) . \tag{9.3.22}$$

We therefore find that the amount of power is greatly enhanced on the tail compared to the naive Gaussian approximation.

Eigenvalue problem. The eigenvalue problem is analogous to the equation for $\chi_{\mathcal{N}}(t,\phi)$; there is no general solution. In the almost constant approximation,

$$\Phi_n''(\phi) - \frac{\alpha}{M_{\rm pl}v_0}\Phi_n'(\phi) + \frac{\Lambda_n}{v_0 M_{\rm pl}^2}\Phi_n(\phi) = 0\,, \tag{9.3.23}$$

the solution reads

$$\begin{aligned}\Phi_n(\phi) = \exp\left(\frac{\alpha\phi}{2M_{\rm pl}v_0}\right)\Bigg[&\alpha_n \exp\left(i\sqrt{v_0\Lambda_n - \frac{\alpha^2}{4}}\frac{\phi}{v_0 M_{\rm pl}}\right)\\ &+ \beta_n \exp\left(-i\sqrt{v_0\Lambda_n - \frac{\alpha^2}{4}}\frac{\phi}{v_0 M_{\rm pl}}\right)\Bigg]\end{aligned} \tag{9.3.24}$$

The first boundary condition imposes that $\beta_n = -\alpha_n$ and the second one implies the trascendental equation (9.3.18). In the almost constant approximation, the eigenvalues Λ_n correspond to (9.3.19).

The eigenfunctions are thus given by

$$\Phi_n(\phi) = \exp\left(\frac{\alpha\phi}{2M_{\rm pl}v_0}\right)\sin\left[\pi\left(n+\frac{1}{2}\right)\frac{\phi}{\phi_{\rm UV}}\right], \tag{9.3.25}$$

which boil down to Eq. (9.3.9) if $\alpha = 0$. Note that the constant slope introduces an exponential suppression of the amplitude in comparison with the flat potential (9.3.7). The coefficients in the expansion (9.0.7), α_n, can be determined as for the flat potential. One first notes that the functions Φ_n of Eq. (9.3.25) form an orthogonal set, in the (extended) sense that

$$\left\langle \Phi_n^{(-\alpha)}, \Phi_m^{(\alpha)} \right\rangle \equiv \int_0^{\phi_{\rm UV}} \Phi_n^{(-\alpha)}(\phi)\,\Phi_m^{(\alpha)}(\phi)\,{\rm d}\phi = \frac{\phi_{\rm UV}}{2}\delta_{n,m}\,. \tag{9.3.26}$$

Making use of the expansion (9.0.7), this gives rise to

$$\left\langle \Phi_n^{(-\alpha)}, P_\phi(\mathcal{N}) \right\rangle = \alpha_n \frac{\phi_{\rm UV}}{2} e^{-\Lambda_n \mathcal{N}}\,. \tag{9.3.27}$$

This gives rise to

$$\begin{aligned}\int_0^{\phi_{\rm UV}} &\Phi_n^{(-\alpha)}(\phi)\,\chi_{\mathcal{N}}(t,\phi)\,{\rm d}\phi = \\ &\frac{\phi_{\rm UV}}{2\mu^2}\frac{1}{\Lambda_n^{(0)} - it}\left[\pi(2n+1) - \frac{2\sqrt{t}\gamma\mu\alpha(-1)^n}{2\gamma\sqrt{tv_0}\cosh\left(\mu\gamma\sqrt{t}\right) + \alpha\sinh\left(\mu\gamma\sqrt{t}\right)}\right].\end{aligned} \tag{9.3.28}$$

In order to Fourier transform this, one again needs to look at the poles. When $\alpha = 0$, there is one pole at $t = -i\Lambda_n^{(0)}$, but the residue at this pole exactly vanishes when $\alpha \neq 0$, and the new pole is located at $t = -i\Lambda_n$. At leading order in v_0, $\Lambda_n \simeq \Lambda_n^{(0)}$, so the second term features a pole at $\Lambda_n^{(0)}$ but this pole is now of order two, so that it does not contribute to the residue theorem. One then ends up with the same expression for α_n as before, namely $\alpha_n = \frac{2\pi}{\mu^2}\left(n+\frac{1}{2}\right)$.

Altogether, in the almost constant approximation the PDF is given by

$$P_\phi(\mathcal{N}) = \sum_n \frac{2\pi}{\mu^2}\left(n+\frac{1}{2}\right)\exp\left(\frac{\alpha\phi}{2M_{\rm pl}v_0}\right)\sin\left[\pi\left(n+\frac{1}{2}\right)\frac{\phi}{\phi_{\rm UV}}\right]e^{-\Lambda_n\mathcal{N}}\,. \tag{9.3.29}$$

Recall that this approximation breaks at $\phi_{\rm uv} \sim 1/\alpha$. A slightly better approximation in this limit can be obtained performing a WKB approximation of Eq. (9.2.5), which can be applied to any potential $v(\phi)$. The procedure begins by absorbing the friction term making a field redefinition

$$\Phi_n = e^{\frac{1}{2}\int \frac{v_\phi}{v^2}d\phi}\tilde{\Phi}_n\,. \tag{9.3.30}$$

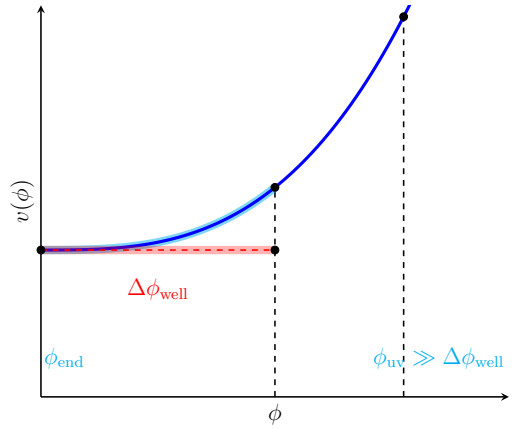

Figure 9.6. Schematic representation of a cacuum dominated potentials with polynomial slope studied in section 9.3. The tail of the PDF can be approximated from a flat potential in the diffusion dominated region $\Delta\phi_{\rm well}$.

In this way, (9.2.5) becomes

$$\tilde{\Phi}_n'' + \left(\frac{\Lambda_n}{v} - \frac{1}{4}\left(\frac{v_\phi}{v^2}\right)^2 + \frac{1}{2}\left(\frac{v_\phi}{v^2}\right)'\right)\tilde{\Phi}_n = 0\,, \tag{9.3.31}$$

which can be solved in plane waves. Specifically, the solution reads

$$\Phi_n = e^{\frac{1}{2}\int \frac{v_\phi}{v^2}d\phi}\left(\alpha_n e^{i\int\theta_n d\phi} + \beta_n e^{-i\int\theta_n d\phi}\right)\,, \tag{9.3.32}$$

where the phase is

$$\theta_n^2 = \frac{\Lambda_n}{v} - \frac{1}{4}\left(\frac{v_\phi}{v^2}\right)^2 + \frac{1}{2}\left(\frac{v_\phi}{v^2}\right)'\,. \tag{9.3.33}$$

For the approximation to hold one requires $\theta_n'/\theta_n^2 \ll 1$. The first initial condition imposes

$$\alpha_n = -\beta_n\,, \tag{9.3.34}$$

so that

$$\Phi_n = 2i\alpha_n\, e^{\frac{1}{2}\int \frac{v_\phi}{v^2}d\phi}\sin\left[\int\theta_n d\phi\right]\,. \tag{9.3.35}$$

The second one determines a transcendental equation for Λ_n,

$$\tan\left[\int_{\phi_{\rm end}}^{\phi_{\rm uv}}\theta_n(\phi)d\phi\right] = -2\frac{v^2(\phi_{\rm uv})}{v_\phi(\phi_{\rm uv})}\theta_n(\phi_{\rm uv})\,. \tag{9.3.36}$$

This is the equation that we need to solve to obtain Λ_n.

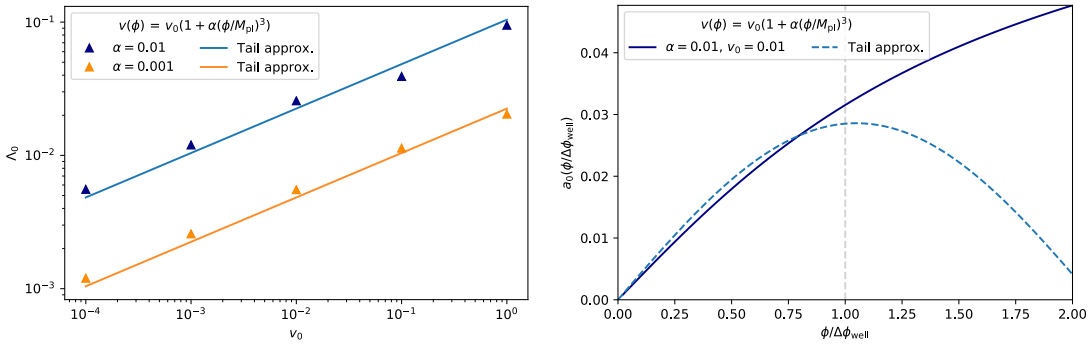

Figure 9.7. Tail of the PDF for a cubic potential: numerical v.s. quantum well approximation. On the left we plot the leading decay Λ_0 as a function of v_0 for different α. On the right we present the amplitude a_0 as a function of ϕ. The approximation only works in the quantum dominated region $\Delta\phi_{\rm well}$.

Vacuum dominated potentials with polynomial slope

Let us now move to more realistic scenarios. We are going to study vacuum dominated potentials with a polynomial slope. In particular, we will focus on the cubic potential

$$v\left(\phi\right) = v_0 \left(1 + \alpha \left(\frac{\phi}{M_{\rm pl}}\right)^3\right) , \tag{9.3.37}$$

although our conclusions can be generalized to other powers. We will evolve the field from $\phi_{\rm uv}$ to $\phi_{\rm end}$. One of the results of the constant and linear potentials was that the tail of the PDF is sensitive to the range $\Delta\phi = \phi_{\rm uv} - \phi_{\rm end}$. However, we do not expect our physical observables to depend on the boundary conditions at high energies. For that reason, we will choose a large enough $\phi_{\rm uv}$ so that the PDF becomes insensitive to its value. In practice, taking $\phi_{\rm uv} \gtrsim 10/\alpha^{1/3}$ is sufficient. Once we are in this regime, the decay of the tail of the PDF becomes a universal prediction of each potential.

Nevertheless, if we have $\phi_{\rm uv} > 1/\alpha^{1/3}$, we cannot use the almost constant approximation applied to the linear case. Still, one could apply a different strategy. Given that the effects of quantum diffusion are dominant only when the potential is very flat, we could delimite this region and approximate it by a flat potential. We could use the "classicality criterion" proposed in [478] stating that stochastic effects dominate only when

$$\frac{v'' \, v^2}{v'^2} \gtrsim 1 \, . \tag{9.3.38}$$

For the cubic potential (9.3.37), this implies that quantum diffusion is relevant in the region

$$\frac{\phi}{M_{\rm pl}} \lesssim \left(\frac{2}{3}\frac{v_0}{\alpha}\right)^{1/3} \sim \Delta\phi_{\rm well} \, . \tag{9.3.39}$$

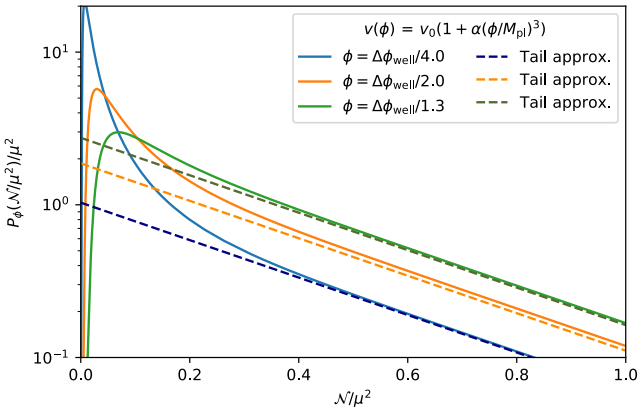

Figure 9.8. Probability distribution function of the number of e-folds $\mathcal{N}$ starting at different initial field values ϕ for a cubic potential. We compare the numerical PDF (solid) with the leading term of the tail expansion (dashed) of a flat potential in the quantum dominated region $\Delta\phi_{\rm well}$. We rescale $\mathcal{N}$ by $\mu^2 = \Delta\phi^2_{\rm well}/v_0$.

Therefore, we could approximate the potential as having only stochastic corrections in this region where we can approximate the potential by a quantum well of duration $\Delta\phi_{\rm well}$. In Fig. 9.6 we schematically represent this situation, although in practice one has $\Delta\phi_{\rm well} \ll \phi_{\rm uv}$.

The nice feature of this *quantum well approximation* is that we can use the analytical formulas for the flat potential obtained in section 9.3. We only need to introduce the value of $\Delta\phi_{\rm well}$ obtained in (9.3.39). For example, the poles of the characteristic function will be located at

$$\Lambda_n = \frac{3^{2/3}\pi^2}{2^{2/3}} v_0^{1/3}\alpha^{1/3}\left(n+\frac{1}{2}\right)^2 . \tag{9.3.40}$$

We have tested this approximation against the numerical results. In Fig. 9.7 we analyze the leading term of the exponential expansion: its exponent Λ_0 and its amplitude $a_0(\phi)$. In the left panel we plot the values of Λ_0 as a function of v_0 for different values of α. The tail approximation lines corresponds to the flat potential with $\Delta\phi_{\rm well} = 1.2\left(\frac{2}{3}\frac{v_0}{\alpha}\right)^{1/3}$. The nice fit with the numerical results confirms that the approximation is valid. On the right plot we present the amplitude a_0 as a function of the initial field value ϕ. As expected, the approximation works nicely but only in the region within $\Delta\phi_{\rm well}$. Finally, we plot together in Fig. 9.8 the full numerical PDF and the leading term of the tail expansion using the quantum well approximation. It can be observed that there is good agreement.

9.4 PBH from exponential tails

Altogether, we have seen that quantum diffusion produces the tail of the PDF to decay exponentially with the number of e-folds for several toy models. As anticipated, this

introduces important differences with the standard picture of a Gaussian PDF. To go from the number of e-folds to the coarse grained curvature perturbations, we can use the relation [478]

$$\zeta_{\rm cg} = \mathcal{N} - \langle \mathcal{N} \rangle \,, \tag{9.4.1}$$

where $\langle \mathcal{N} \rangle$ is the mean number of e-folds, which can be obtained solving a differential equation

$$\langle \mathcal{N} \rangle'' - \frac{v_\phi}{v^2} \langle \mathcal{N} \rangle' + \frac{1}{v M_{\rm pl}^2} = 0 \tag{9.4.2}$$

with boundary conditions $\langle \mathcal{N} \rangle(\phi_{\rm end}) = \langle \mathcal{N} \rangle'(\phi_{\rm uv}) = 0$.[1] Subsequently, we only need to integrate the PDF to obtain the abundance of PBH via (6.0.1) to find

$$\beta(M) = \sum_n \frac{1}{\Lambda_n} \, a_n(\phi) \, e^{-\Lambda_n (\zeta_c - \langle \mathcal{N} \rangle)} \,, \tag{9.4.3}$$

where we have used the tail expansion (9.0.7). Accounting for a particular inflationary model just implies choosing the particular amplitudes $a_n(\phi)$ and exponents Λ_n.

The first thing to notice from the above expression is that the exponential decay of the tail of the PDF translates into an exponential dependence with the threshold ζ_c. This expression can be contrasted with the Gaussian result (6.0.4), which in the limit of small fluctuations $\sigma \ll \zeta_c$ becomes

$$\beta_{\rm G}(M) = \frac{\sigma}{\sqrt{2\pi}\zeta_c} \, e^{-\frac{\zeta_c^2}{2\sigma^2}} \,. \tag{9.4.4}$$

Clearly, the Gaussian approximation will be many orders of magnitude suppressed compared to the result from stochastic inflation. This difference was first noted for a flat potential in [474]. Here, we have extended that result showing that more general potentials have the same type of exponential tail.

A second point to stress is that the poles Λ_n of the characteristic function turn out to be the relevant quantity to asses the abundance of PBH. Moreover, in the case of polynomial potentials these poles are a universal prediction of each potential, as we have explicitly shown for the cubic potential in (9.3.40). This is an interesting property that can serve identifying models with more or less PBH production.

Finally, let as mention that there are other inflationary models that also display this χ-square statistics for the PDF. For instance, models of axion inflation in which the gauge field source the curvature perturbations display this same behavior [481]. The implications of a χ-square distribution for the GWs signatures have been studied in detail in Ref. [457].

[1] This equation follows directly form the definition of the characteristic function (9.1.1) and the differential equation (9.1.3) it satisfies.

Part IV

Conclusions

CHAPTER

10

Executive summary

Gravitational wave astronomy has opened a new channel to listen to the cosmos. This offers novel opportunities to probe the pillars of the standard cosmological model. In this thesis we have focused on the dark sectors of the universe. We have proposed and applied tests of dark energy and General Relativity with present and future GW detectors. Moreover, we have investigated how to properly compute the abundance of black-holes formed in the early universe, which have strong implications on their contribution to the dark matter and on their GW signatures. In the following we summarize the main results we have obtained.

10.1 Dark energy in light of gravitational wave astronomy

Although the present accelerated expansion of the universe is a robust pillar of modern cosmology, its origin remains puzzling. Beyond the cosmological constant paradigm, there is a plethora of DE proposals, many of them extending Einstein's theory of gravity. Reducing the number of contenders and probing gravity at large-scales are thus priorities. Tests of the propagation of GWs are especially powerful for this task because they are clean (compared to effects in the emission of GWs) and precise (since small modification accumulate over large travel distances). We have explored three main probes of the GW propagation: the speed of GW, the GW luminosity distance and GW oscillations, covered in chapter 3, 4 and 5 respectively.

The speed of GWs and DE after GW170817

One of the most fundamental properties of a wave is the speed at which it propagates. GR predicts that GWs propagate at the speed of light c. Nonetheless, in DE scenarios beyond GR, the dark fluid responsible for the cosmic expansion could act as a medium affecting the GW velocity. We have investigated how to compute the propagation speed when GWs travel across general backgrounds in extensions of Einstein's gravity. We identified the effective gravitational wave metric $\mathcal{G}^{\mu\nu}$ as the relevant quantity determining when an anomalous speed arise. Non-luminal speeds will appear whenever $\mathcal{G}^{\mu\nu}$ is not conformally related to the original metric $g^{\mu\nu}$ (cf. Fig. 3.1). In the realm of scalar-tensor gravity, this allowed us so set the two conditions for an anomalous GW speed: i) a non-trivial scalar field configuration and ii) a derivative coupling to the curvature introducing a Weyl tensor in the equations of motion. This criterion provides a clear-cut way to distinguish

two classes of gravitational theories, those for which the speed of GWs is *exactly equal* to the speed of light, and those in which it can vary depending on the theory parameters and the background configuration of the scalar field. We used this criterion to predict which gravity theories would be affected by a measurement of the speed of GWs.

An anomalous GW speed can appear in different situations. In cosmology, in order to explain DE, one typically demands that the scalar field background varies with the Hubble rate, $\dot{\phi} \sim H_0$. As we saw in section 3.2, this is enough to induce $\mathcal{O}(1)$ deviations from the speed of light in theories like covariant Galileons. In fact, for this particular theory, it turned out that the only sector of the theory that does not modify the speed of GWs is ruled out by other cosmological tests. But an anomalous speed does not only appear due to the cosmological evolution. It can also be sourced by the spatial profile of the scalar field as we have shown in section 3.3. This happens for instance in gravity theories where BHs can have scalar hair such as scalar Gauss-Bonnet. Similarly, the screening mechanism preventing small scale modifications of gravity also introduces scalar gradients modifying the speed of GWs. However, the size of the deviation from the speed of light depends on the system under consideration.

Everything changed after GW170817, a binary neutron star merger followed by a short gamma-ray burst GRB170817A just 1.7s later. This was the birth of multi-messenger GW astronomy but also implied the death of many DE models. The (almost) coincident arrival of EM and GW signals places one of the strongest bounds available on a large class of scalar-tensor theories that predict an anomalous GW speed. The severe constraints on Galileons extends to other scalar-tensor theories: without fine tuning, the quartic and quintic sector of Horndeski, as well as GLPV and several other beyond Horndeski Lagrangians are effectively ruled out as dark energy or late universe modifications of gravity. These theoretical classes include some interesting models, such as accelerating solutions due to the weakening of the gravitational force [482] and self-tuning theories that attempt to solve the cosmological constant problem, and which rely on non-minimal derivative couplings to curvature [93].

Despite the strong bounds, theories remain that avoid this constraint and thus can still be used to explain DE (see Fig. 3.4). Within Horndeski's theory these include only the simplest modifications of gravity. Beyond Horndeski theory, viable gravities can be obtained in two ways. One can apply a derivative-dependent conformal transformation to those Horndeski models with $c_g = 1$, since it does not affect their causal structure. Alternatively, one can implement a disformal transformation, which does alter the GW-cone, designed to precisely compensate the original anomalous speed of the theory. However, fine-tuning the cosmological evolution is not a viable solution since any perturbation will re-introduce an anomalous speed (note that $|c_g/c - 1| \leqslant 10^{-15}$).

The constraints of GW 170817 extends further into the landscape of gravity theories (see Fig. 1.2). In the case of vector-tensor and scalar-vector-tensor theories, there are several couplings to the curvature that now will be extremely constrained because they modify the speed of GWs, e.g. $R_{\mu\nu}v^\mu v^\nu$ in vector DE [360]. In particular, this test has an impact on Einstein-Aether theories, including some sectors of Hořava gravity, and more general frameworks such as Generarlized Proca theories. TeVeS and MOND-like theories

are as well critically affected by this bound. Massive gravity, bigravity and multi-gravity remain viable as long as the graviton mass is small and matter couples minimally to one of the metrics (see section 1.1.1 for a review on the different modifications of gravity).

Testing gravity with the GW luminosity distance

Another fundamental probe of the propagation of GWs is their amplitude, which scales with the inverse of the distance from the source to the observer. GR predicts that the GW luminosity distance $d_L^{\rm gw}$ is equal to the EM luminosity distance $d_L^{\rm em}$, which over cosmological backgrounds is just an integral of the Hubble length (4.0.2). Modified gravity theories generically alter this prediction. The ratio $d_L^{\rm gw}/d_L^{\rm em}$ thus become a powerful discriminator of any additional friction term along the propagation.

We have determined how $d_L^{\rm gw}$ would differ from $d_L^{\rm em}$ in generic gravity theories when there is a modification of the absorption of the cosmic medium and an anomalous GW speed. Our main result is encapsulated in equation (4.1.17). When $c_g = c$, the ratio $d_L^{\rm gw}/d_L^{\rm em}$ still is not one. This suggests a promising avenue to probe those DE models that survived GW170817.

The GW luminosity distance can be tested with GWs alone, using number counts, or with multi-messenger events, using either standard or dark sirens. We have contributed to forecast the prospects of testing a modified $d_L^{\rm gw}$ with LISA standard sirens. These sources are particularly interesting because they are expected to be detected up to redshift $z \sim 5$. Therefore, LISA has significantly more constraining power than present ground-based detectors (see Fig. 4.1). Using updated estimates for LISA configuration and sensitivity curve, we found that $d_L^{\rm gw}/d_L^{\rm em}$ could be bound at the percent level ($1-4\%$ depending on the assumptions on the mock catalogue of massive BH binaries).

GW oscillations as an indicator of additional cosmological fields

Beyond the cosmological constant paradigm, dark energy models typically introduce additional cosmological fields. These extra modes can also be probed with the propagation of GWs. Over homogeneous and isotropic cosmological backgrounds, tensor perturbations only couple to other tensor modes at linear level. A paradigmatic example of a theory with two types of tensor perturbations is bigravity. Nonetheless, theses extra tensor modes could be the effective configuration of other types of fields, for instance gauge fields. In analogy to neutrino oscillations, if there are tensor modes interacting while propagating, there could be gravitational wave oscillations.

We have developed a method to study GW oscillations in full generality. We divide the possible interactions in mass, friction, velocity and chiral mixing. All of them can produce oscillations in the amplitude of the GWs but their relative relevance is different. In each case we identify the frequency of oscillation and compare the WKB and large-k approximation schemes. Interestingly, we find that although the tensor perturbations coupled to matter propagate at the speed of light, if the the other tensor has an anomalous speed and there are GW oscillations, the effective propagation speed of the GWs can differ from the speed of light.

GW oscillations leave several distinctive signals that we have studied in detail. Assuming that the production of GWs is not modified, the mixing of the tensor modes along the propagation translates into a time dependent deficit of the original GW. For a given source, this mixing is fixed by the distance and would make the GW strain to be globally reduced. If the mixing depends on the frequency of the wave, this would mean a time dependent modulation of the wave-form. These two effects happen in models with friction and mass mixing respectively, as exemplified in Fig. 5.8. But if the signal is dimmer, this would be interpreted as the source being further apart. Therefore, GW oscillations also modify the GW luminosity distance in an oscillatory manner (see Fig. 5.9). This characteristic pattern can be probed with multiple standard sirens at different redshifts. Note also that in the case of having a complete conversion from one mode to the other, this would imply that no GWs would be detected at certain redshifts. An effect that is easy to distinguish from the astrophysical modeling. Moreover, if there is an induced anomalous speed this can be probed with multi-messenger events like GW170817. Finally, with a network of detectors one could in principle detect also a polarization dependent GW oscillation.

We have applied our general formalism to demonstrate the improved capability of LISA to probe GW oscillations in models with extra tensor interactions such as bigravity. We compute predictions for bigravity using a high frequency expansion, focusing on the high mass regime $m_g \gg H_0$. In this region of parameter space, GW oscillations occur in the mHz range and the theory is not generically able to account for cosmic acceleration. Standard sirens at cosmological distances, as will be observed by LISA, have the potential to constrain the mass range $m_g \gtrsim 10^{-25}$eV for most mixing angles, 3 orders of magnitude better than current LIGO-Virgo detections. We expect these constraints to be similar in other theories, for instance with a friction mixing. Moreover, we also anticipate that including frequency-dependent effects on the waveform will improve these bounds.

10.2 Primordial black-hole probes of the early universe

Jumping back on time, another pillar of the standard cosmological model are primordial density perturbations becoming seeds of the structure of the universe. Still, our knowledge of them is quite limited. In terms of the primordial power spectrum (PS), only the largest scales have been constrained so far, mostly by the Cosmic Microwave Background (CMB). Thus, observations allow large, small scale fluctuations generated in the early universe. If large enough (possibly of order 10^{-1} compared to the CMB's 10^{-5}), such fluctuations would have collapsed upon horizon re-entry to form primordial black-holes (PBHs). Therefore, if PBHs were to exist, they would constitute a direct, new window to unproved periods of the early universe. Conversely, if they are not found, they can be used to place bounds on different high-energy scenarios. As a bonus, constraining the population of PBHs also teach us about how much they can contribute to the DM budget. Among other methods, PBHs can be constrained by the GWs they emit when binaries merge. We have studied the formation of PBHs in single field models of inflation in chapter 6 and proposed a critical Higgs inflation scenario in chapter 7. We have also

determined the relevance of quantum diffusion for the formation of PBHs developing an iterative method and computing the tail of the curvature perturbation distribution in chapters 8 and 9 respectively.

PBH formation in single field inflation

One way of producing PBHs it to enhance the primordial PS of curvature perturbations ζ. In single field inflation this can be achieved having a second plateau in the potential. In this manner, the inflaton slows down while crossing the inflection point and the curvature perturbations grow ($\zeta = \delta\phi/\sqrt{2\epsilon}$ in the uniform-curvature gauge). However, this rapid deceleration may produce a violation of slow-roll due to $|\epsilon_2| > 1$.

We have reanalyzed the enhancement of the power spectrum in inflationary models with a quasi-inflection point. Beyond SR, the PS could grow either at sub-horizon or super-horizon scales. At sub-horizon scales, the PS grows when the field slows down since $\mathcal{P}_\zeta \sim V(\phi)/\epsilon + \mathcal{O}(\epsilon^2)$. At super-horizon scales, there is a constant and an evolving mode, which amplifies or freezes depending on the sign of the friction. Whenever $2 - \epsilon + \epsilon_2 < 0$, the evolving mode exponentially grows, which can only be achieved if there is a violation of SR. Thus, during this period outside of SR, the PS is also enhanced. Both mechanisms serve to enlarge the production of PBHs. Depending on the time in which the inflaton enters the plateau and on how long it takes to cross it, the spectrum of PBHs will peak at different masses with a different width.

A particle physics motivated scenario: critical Higgs Inflation

We have proposed a particle physics motivated model in which PBHs form from a quasi-inflection point in the potential. In this critical Higgs inflation scenario, the Standard Model Higgs non-minimally coupled to gravity acts as the inflaton in the early universe. Taking into account the renormalization group equation (RGE) running of both the Higgs self-coupling λ and the non-minimal coupling to gravity ξ, we find regions of parameter space allowed by the Standard Model for which the inflaton-Higgs potential acquires a second plateau at smaller scales, around the critical point in the RGE flow $\lambda(\mu) \simeq \beta_\lambda(\mu) = 0$. This plateau gives an ultra-slow-roll evolution of the Higgs, inducing a high and broad peak in the curvature power spectrum.

We have explored both the early and late universe observables of this model. Regarding the CMB, we find regions in the parameter space compatible with Planck 2018 and displaying a large enhancement of the perturbations at smaller scales. However, larger and broader peaks tend to have a worse agreement since the spectral index becomes smaller. Interestingly, this model predicts a tensor-to-scalar ratio within the reach of present and next-generation B-mode experiments, $r \sim 0.04$. In terms of the population of PBHs, this scenario predicts an approximate lognormal distribution of masses. Their evolution depends on the merger and accretion, which is subject to the initial distribution and feedback with baryonic matter. Under the assumption of a highly clustered initial distribution, we find that PBH could constitute a fraction of the dark matter with masses ranging from 0.01 to $100 M_\odot$.

Critical Higgs inflation and similar PBH scenarios can be tested with GWs in several ways. A golden probe would be to detect a sub-solar BH, since there is not known astrophysical mechanism that could form them. In this type of model, $0.1 - 1\%$ of the binaries would involve a BH with a mass below the Chandrasekhar limit. Also, accumulating binary BH merger detections one can constrain the mass distribution. Probably more promising are measurements of the effective spin, which PBH models generically predict to be close to zero, or merger rates at high-redshift, in which astrophysical BHs would not have time to form. Finally, PBHs scenarios like critical Higgs inflation produce stochastic backgrounds both at formation, by second order perturbations, and at merger, from the unresolved binaries.

Important implications of quantum diffusion

PBHs are generated from large curvature perturbations. However, perturbative techniques applied to inflation neglect non-linear contributions that can be relevant when $\zeta \sim 1$. For this reason, it is more appropriate to study PBH using an effective field theory for super-horizon perturbations in which short-wavelength modes behave as a classical stochastic noise. This is nothing but stochastic inflation.

We have investigated PBH production in single field models of inflation with a quasi-inflection point taking into account the back-reaction of quantum fluctuations. Since the production of PBHs is related to a deviation of SR, we have considered the system of Langevin equations associated to both the coarse-grained field $\bar{\phi}$ and momentum $\bar{\pi}_\phi$ in the Hamilton-Jacobi formalism of stochastic inflation. In order to find the multi-variate probability density function associated to this stochastic problem, we have solved the system of first-order differential equations for the statistical moments $\langle \delta\phi^m \delta\pi^n \rangle$. With these moments one can construct the characteristic function $\chi(u_\phi, u_\pi)$ that can be used to compute the probability density function (PDF). The methodology that we have developed can be applied to study quantum diffusion effects beyond slow-roll in general. We have particularized its use to find the probability that a given fluctuation is above a certain threshold $P(\zeta > \zeta_c)$, which is the relevant question to know if PBHs are formed.

We have analyzed the implications of quantum diffusion beyond SR for the production of PBHs. Inflationary models with a quasi-inflection point induce an exponential growth of the curvature modes whenever $3 - \epsilon + \epsilon_2 < 0$. We find that this enhancement of the quantum fluctuations also enlarges the stochastic noise. In turn, this produces a peak in the power spectrum on top of the spectrum of ζ_k without diffusion. Such a diffusion peak has a clear effect amplifying the PBH production. Moreover, we observe that this large stochastic noise also induce important non-Gaussian contributions. These NG contributions could either enlarge or decrease the PBH production. For our toy model potential, we realize that the dominant contribution is given by the fourth moment. It introduces a positive contribution that augments the weight of the tails of the PDF, leading to a significantly higher probability to form PBHs. Adding both contributions, we find that for our case of study quantum diffusion substantially boosts the formation of PBHs.

Our results show that quantum diffusion beyond SR can have an important role in

determining the abundance of PBHs for single-field models with a quasi-inflection point. These effects can be parametrized with the dynamics of the Hubble-flow parameters $\epsilon_n(N)$. We expect that our conclusions could be extended to different potentials studied in the literature that share a similar behavior of ϵ_n. We believe that the predictions for the spectrum of PBHs in this class of models should be revisited in the future accounting for quantum diffusion.

The role of the tail of the PDF

The abundance of PBH is given by the probability that a curvature perturbation is above a certain threshold. Therefore, the production of PBHs is controlled by the tail of the PDF $P(\zeta)$. We have seen that non-Gaussian corrections induced by the stochastic noise can be relevant in the inflection point. It would then be relevant to have non-perturbative techniques to fully characterize the tail without relying on an expansion in the statistical moments.

We have developed novel techniques to compute the tail of the PDF in the context of SR inflation. We demonstrate that the exponential decay of the tail is controlled by the poles of the characteristic function. These poles can be computed numerically or analytically in certain cases, such as the flat and the almost constant linear potential that we study. Alternatively, one could obtain the tail of the PDF from the eigenvalues of an equivalent eigen problem derived from the Fokker-Planck equation. We find that the poles and eigenvalues depend on the parameters of the potential and on the boundary conditions. Namely, the range of possible initial field values considered from $\phi_{\rm uv}$ to $\phi_{\rm end}$.

We also investigate more realistic vacuum dominated potentials with a polynomial slope. There, we are interested in the regime in which the tail of the PDF is not sensitive to the boundary conditions at high energies. We find that the tail of the PDF derived from the stochastic evolution though the entire potential can be very well approximated by a flat potential in the region where quantum diffusion dominates, determined by $v^2v''/v'^2 \gtrsim 1$. We compare the numerical results for the poles with this quantum well approximation obtaining nice agreement. This implies that the tail of the PDF is a universal property for these potentials.

We conclude that the tail of the PDF is significantly more relevant than the usual Gaussian estimate. This is because quantum diffusion induces an exponential decay linear to the curvature $\sim e^{-\Lambda_n\zeta}$ rather than quadratic $\sim e^{-\zeta^2/2\sigma^2}$. Of course, this also implies that for a given threshold ζ_c, the amount of PBH formed is many orders of magnitude larger than the Gaussian prediction. This has important implications for inflationary model building and PBH phenomenology.

CHAPTER

11

Future prospects

In the last four years we have witnessed a vibrant beginning of GW astronomy: from the excitement of the first detection, to the routine of the following events and back to the euphoria of a multi-messenger historic achievement. Binary BH (BBH) mergers have taught us about a (unexpected) population of heavy BHs (see left panel of Fig. 1.4) as well as their signals have served to strongly constrain the mass of the graviton (see section 2.5). More notoriously, GW170817, a binary neutron star (NS) event, brought the first standard siren measurement of the Hubble constant (see section 2.2.1) and an impressive bound on the speed of GWs, $|c_g/c - 1| \leqslant 10^{-15}$, that swept out many DE contenders (see section 3.4.1).

The following years are no less exciting, with observational improvements pointing in three main directions: number of detections, sky localization and distance range (see Fig. 2.7). However, these observational opportunities will also pose theoretical challenges, as we will be crossing both discovery and precision frontiers. In what comes next, we overview these challenges and opportunities, focusing on those related to unveil the nature of dark energy and to discriminate the origin of the observed BHs.

Before that, it is important to remember that GW observatories will not be alone probing the dark universe. Cosmological surveys such as LSST, Euclid or WFIRST will have a fundamental role (see recent reviews of their science cases in [483, 484], [485] and [486] respectively). With respect to the theoretical challenges, they will constrain among others the DE equation of state and possible deviations of GR in the growth of structure [483, 485, 486]; as well as the DM distribution on large scales [484, 485]. These complementary observations could be useful in breaking degeneracies. For example, we have seen here how the constraint on the speed of GWs together with the integrated Sachs Wolfe effect ruled out covariant Galileons as a DE contender (see section 3.2). From the observational side, a key task of future surveys will be to deliver complete and uniform galaxy catalogs. This is essential for instance in order to properly identify the redshift of a GW signal with the statistical method (see section 2.2.1). We expect that the most challenging questions about DE and DM will only be addressed with a powerful synergy between GW astronomy and other cosmological probes.

11.1 Theoretical challenges

Arguably, there are many theoretical challenges regarding GW astronomy and fundamental physics. Probably two are the main issues: *i)* develop novel formalisms to cover unexplored regimes and *ii)* understand degeneracies with astrophysics. In this thesis we have studied how to probe dark energy models with the propagation of GWs and how to learn about the early universe with primordial BHs. We now examine some unexplored regimes that could be investigated using or extending the results of this thesis.

GW propagation on general space-times.

Just as electromagnetic (EM) waves can scan materials, GWs can probe the medium in which they propagate. We have seen throughout this thesis that the cosmological propagation of GWs provides powerful tests of the properties of DE and the content of the universe (see part II). However, modification in the propagation are not restricted to cosmological backgrounds. For instance, we have already anticipated that in regions in which there is a spatial gradient of the scalar field an anomalous GW speed could be induced (see section 3.3). Examples of this are screened regions in which the scalar field develops a background that prevents fifth forces to appear in order to pass local tests of gravity. In fact, we generically expect GWs to cross several screened regions in their travel from the source to the observer. The key point will be to understand the possible interplay of additional polarizations. In particular, if there could be a mixing among them exciting polarizations that were not originally produced (in an analogous way to the GW oscillations that we studied in chapter 5). If those polarizations couple to matter, they could be detected with a network of ground-based detectors. This would be a clear smoking gun of a modification of gravity.

Frequency dependent modifications of the propagation of GWs.

Modifications of the propagation of GWs depend on space and time but may also depend on the frequency of the wave. The best known example of this is having a graviton mass, which modifies the dispersion relation (recall (3.0.2)). The advantage of this type of modification is that they can be probed with GWs alone and on single events. This is simply because the chirping of the GW signal makes its frequency to vary with time. For instance, we have shown that GW oscillations with a mass mixing leave a distinctive oscillatory pattern in the wave-form (see Fig. 5.8). Since GW detectors are very sensitive to frequency dependent modifications, it would be interesting to search further for this kind of effects in the landscape of modified gravity. For instance, it has been shown recently that if there is a breaking of the EFT at the frequencies of GW detectors, the dispersion relation is modified [389]. Similarly, it has been advocated in [487] that the possible decay of GWs into DE could imprint also frequency dependent friction terms and dispersion relation.

Curvature perturbation distribution tail beyond slow-roll.

PBHs can be copiously produced in single field models of inflation with a quasi-inflection point (see part III). Quantum diffusion dominates the inflationary dynamics around

this point, where the slow-roll approximation is violated. In chapter 8 we developed a formalism to study the impact of the quantum back-reaction through the stochastic noise. To asses how the classical picture was modified, we devised an iterative process to solve the statistical moments of the probability density function (PDF) of the curvature perturbations. We found that non-Gaussian corrections were indeed relevant. In parallel, we found new techniques to compute directly the tail of the PDF in chapter 9, which as a first step we applied to SR inflation. Therefore, a natural extension of this work would be to generalize the methods to compute the tail of the PDF to inflationary models beyond SR. Having an accurate description of the tail of the PDF is crucial to properly infer the properties of the PBH population.

PBH formation from non-Gaussian fluctuations and their evolution.

Examining GWs from BBH mergers is a direct route to trace back their origin. BH demographics key parameters include mass spectrum, spin distribution, eccentricities and merger rate. In order to reliably predict them in a PBH scenario, it is necessary to compute accurately both their properties at formation and their subsequent evolution (which of course depends on the initial clustering). Present analysis of the formation of PBHs and their initial distribution assume that the curvature perturbations are Gaussian. In chapters 8 and 9 we found however that non-Gaussian corrections are relevant. Accordingly, an interesting follow-up of these results would be to obtain the spatial distribution of PBHs and the probability to form binaries starting with a non-Gaussian PDF. This might have relevant implications on the GW observables and constraints of these PBH scenarios.

11.2 Observational opportunities

The future of gravitational wave astronomy is promising. We are literally at the beginning of a new era and many breakthroughs are yet to come. So far we have started to characterize a large population of binary BHs and we had a first instance of multi-messenger GW astronomy. In the coming years many more detections will allow to perform statistically significant studies but also to detect rare events with very high signal to noise ratio to make precision measurements. Moreover, new frequencies will be opened, namely the mHz from space-based detectors, certainly leading to discoveries as well as allowing for new techniques such us multi-band GW astronomy. Here we highlight what some of these observational opportunities could contribute to the quest of the dark side of the universe.

Second generation detectors

Second generation (2G) detectors still are on their way to arrive at their design sensitivity, as we have summarized in Fig. 2.7. This means that merger rates are expected to increase in the future observational runs. In fact, we are now witnessing one of such steps forward in LIGO/Virgo O3 compared to the previous runs O1-2. Accumulating a large number of detections possibilities statistical analysis. Regarding BH demographics, we expect this

to be the main improvement: a precise measurement of the mass spectrum, effective spin and merger rate. Collecting statistics also has implications for the study of the cosmic expansion, enlarging the number of multi-messenger events to be used as standard sirens or the number of BBHs to be used as dark sirens. This will improve the measurement of the Hubble constant (see section 2.2.1) as well as the constraints on the propagation of GWs (see for example the left panel of Fig. 4.1).

Moreover, all of the planned 2G detectors are not yet online. KAGRA is expected to join by the end of O3 and IndIGO just in the last run (again see Fig. 2.7). Having a large network of detectors is essential to test the number of polarizations, a fundamental property of each gravity theory. This will open new opportunities to test for instance chiral GW oscillations (see section 5.2.4).

Third generation detectors

Third generation (3G) detectors will improve their predecessors in multiple ways [488]: observing more distant binaries, in a wider frequency range and with higher precision (compare the expected sensitivities in Fig. 2.5). All of these improvements will be relevant to unveil the origin of the BBHs. Detecting mergers at high-redshift will serve to determine if BBHs correlate with star formation, which peaks around $z \sim 2$. Enlarging the frequency band will be interesting in both ends. At low frequencies, reaching down to 1Hz, 3G detectors will be sensitive to intermediate mass BHs with masses possibly above $100M_{\odot}$. At high frequencies, reaching up to 5kHz, the possibility of sub-solar mass BHs could be much better constrained. Finally, detecting high signal events implies a better characterization of the source, going beyond the chirp mass and the effective spin to which 2G detectors are basically sensitive to. This is again basic to distinguish a primordial versus an astrophysical origin.

As the sensitivity will improve significantly, 3G detectors could detect many orders of magnitude more events. This will ensure precision measurements of H_0 and the propagation of GWs beyond 2G capabilities. Moreover, it is expected that some signals will be strongly lensed [489]. This could be a perfect arena to test modifications of gravity.

Space-based detectors

A million kilometer interferometer in space could detect GWs in the mHz band. This is the goal of LISA. LISA will open a completely new channel to detect very massive BBHs, galactic binaries and possibly cosmological backgrounds. If identified by a EM counterpart, SMBH binaries would become standard sirens at cosmological scales, $z \sim 2 - 5$. We have demonstrated that LISA could probe GW oscillations thanks to these standard sirens (see Fig. 5.11). In the same manner, it could bound generic modifications of the GW luminosity distance (see right panel of Fig. 4.1). In addition, we have proposed to use LISA verification binaries to measure the speed of GWs at these frequencies, which we detail in appendix B. Lastly, if they exist, LISA would be sensitive to different stochastic backgrounds from PBHs, which could constrain models such as critical Higgs inflation (see section 7.2.2). There could be a background from the unresolved PBH

mergers throughout the cosmic history, but also another background generated at the time of formation of PBHs from the second order perturbations [457].

The reach of GW astronomy can be enhanced if GW events are heard at different frequencies. Heavy BBHs like GW150914 could be detected months in advance by LISA [490]. This will again open new opportunities. For instance, with multi-band events one could combine the information that is better measured in space, such as eccentricities, with the inference on the mass and spin obtained at the merger. This could improve the source characterization. Similarly, anticipating the detection of a signal form space will improve its localization on Earth, making more effective the search for an EM counterpart. At the same time, multi-band events with precise localization are perfect candidates to probe additional polarizations.

Conclusiones

La astronomía de la onda gravitacionales (OG) ha abierto un nuevo canal para escuchar el cosmos. Esto ofrece nuevas oportunidades para testar los pilares del modelo cosmológico estándar. En esta tesis nos hemos centrado en la energía oscura (EO) y la materia oscura (MO). Hemos propuesto y aplicado pruebas de la EO y Relatividad General (RG) con detectores de OG actuales y futuros. Además, hemos investigado cómo calcular adecuadamente la abundancia de agujeros negros (AN) formados en el universo primitivo, que tienen fuertes implicaciones en su contribución a la MO y la producción de OG. A continuación resumimos los principales resultados que hemos obtenido.

Iluminando la energía oscura con las ondas gravitacionales

Aunque la actual expansión acelerada del universo es un pilar básico de la cosmología moderna, su origen sigue siendo desconocido. Más allá del paradigma cosmológico estándar, hay una gran cantidad de propuestas de EO, muchas de ellas extendiendo la teoría de la gravedad de Einstein. Reducir el número de contendientes y explorar la gravedad a grandes escalas son, por lo tanto, prioridades. Las pruebas de propagación de OG son especialmente potentes para esta tarea porque son limpias (en comparación con los efectos en la emisión de OG) y precisas (ya que las pequeñas modificaciones se acumulan en grandes distancias de viaje). Hemos explorado tres pruebas principales de la propagación de OG: la velocidad de las OG, la distancia de luminosidad de las OG y las oscilaciones de OG, cubiertas en los capítulos 3, 4 y 5 respectivamente.

La velocidad de las OG y la EO tras OG170817

Una de las propiedades más fundamentales de una onda es la velocidad a la que se propaga. La RG predice que las OG se propagan a la velocidad de la luz c. No obstante, en los escenarios de EO más allá de la RG, el campo responsable de la expansión cósmica podría actuar como un medio que afecta la velocidad de OG. Hemos investigado cómo calcular la velocidad de propagación cuando las OG viajan a través de medios generales en extensiones de la gravedad de Einstein. Identificamos la métrica de onda gravitacional efectiva $\mathcal{G}^{\mu\nu}$ como la cantidad relevante que determina cuándo suEGR una velocidad anómala. Las velocidades no luminales aparecerán siempre que $\mathcal{G}^{\mu\nu}$ no esté relacionado de forma conforme con la métrica original (ver Fig. 3.1). En el ámbito de la gravedad escalar-tensor, esto nos permitió establecer las dos condiciones para una velocidad de OG anómala: i) una configuración de campo escalar no trivial y ii) un acoplamiento derivativo a la curvatura introduciendo un tensor de Weyl en las ecuaciones del movimiento. Este criterio proporciona una forma clara de distinguir dos clases de teorías gravitacionales, aquellas para las que la velocidad de las OG es *exactamente igual* a la velocidad de la luz, y aquellas en las que puede variar dependiendo de los parámetros de la teoría y la configuración del campo escalar. Utilizamos este criterio para predecir qué teorías de la gravedad se verían afectadas por una medición de la velocidad de las OG.

Una velocidad anómala puede aparecer en diferentes situaciones. En cosmología, para

explicar la EO, uno típicamente exige que el campo escalar varíe con la tasa de Hubble, $\dot{\phi} \sim H_0$. Como vimos en la sección 3.2, esto es suficiente para inducir desviaciones de orden 1 de la velocidad de la luz en teorías como Galileons. De hecho, para esta teoría particular, resultó que el único sector de la teoría que no modifica la velocidad de las OG está descartado por otras pruebas cosmológicas. Pero una velocidad anómala no aparece solo por la evolución cosmológica. También puede obtenerse mediante el perfil espacial del campo escalar, como hemos mostrado en la sección 3.3. Esto sucede, por ejemplo, en las teorías de la gravedad en las que los AN pueden tener "pelo escalar", como en scalar Gauss-Bonnet. De manera similar, el mecanismo de apantallamiento que evita las modificaciones a pequeña escala de la gravedad también introduce gradientes escalares que modifican la velocidad de las OG. Sin embargo, el tamaño de la desviación de la velocidad de la luz depende del sistema que se esté considerando.

Todo cambió después de GW170817, una fusión binaria de estrellas de neutrones seguida de una breve explosión de rayos gamma GRB170817A solo 1,7 s más tarde. Este fue el nacimiento de la astronomía multi-mensajera de OG, pero también implicó la muerte de muchos modelos EO. La llegada casi simultáneas de ondas electromagnéticas y OG implica uno de los límites más fuertes disponibles en una gran clase de teorías de escalar-tensor ue predicen una velocidad de OG anómala. Las severas cotas en Galileons se extienden a otras teorías escalar-tensor: sin ajustes finos, el sector quártico y quíntico de Horndeski, así como GLPV y otros modelos más allá de Horndeski se descartan como EO o modificaciones tardías de la gravedad del universo. Estas clases teóricas incluyen algunos modelos interesantes, como soluciones de aceleración debidas al debilitamiento de la fuerza gravitacional [482] y teorías de autoajuste que intentan resolver el problema de la constante cosmológica y que se basan en acoplamientos derivados no mínimos a la curvatura [93].

A pesar de los fuertes límites, existen teorías que evitan esta restricción y, por lo tanto, aún se pueden utilizar para explicar la EO (ver Fig. 3.4). Dentro de la teoría de Horndeski, éstas incluyen solo las modificaciones más simples de la gravedad. Más allá de la teoría de Horndeski, las gravedades viables se pueden obtener de dos maneras. Se puede aplicar una transformación conforme dependiente del término cinético del escalar a modelos de Horndeski con $c_g = 1$, ya que no afecta su estructura causal. Alternativamente, uno puede implementar una transformación disforme, que altera el cono de OG, diseñado para compensar con precisión la velocidad anómala original de la teoría. Sin embargo, ajustar la evolución cosmológica no es una solución viable, ya que cualquier perturbación reintroducirá una velocidad anómala (tenga en cuenta que $|c_g/c - 1| \leqslant 10^{-15}$).

Las cotas de GW170817 se extienden a través del paisaje de las teorías de la gravedad (ver Fig. 1.2). En el caso de las teorías vector-tensor y escalar-vector-tensor, existen varios acoplamientos a la curvatura que ahora estarán extremadamente limitados porque modifican la velocidad de las OG, por ejemplo. $R_{\mu\nu}v^\mu v^\nu$ en el caso de EO vector [360]. En particular, esta prueba tiene un impacto en las teorías de Einstein-Aether [306], que incluyen algunos sectores de Hořava gravity [45], y marcos más generales como las teorías de Proca generalizadas [491]. TeVeS [126] y las teorías similares a MOND [342,492] también se ven muy afectados por este límite. La gravedad masiva [137], bigravity [150]

y la multi-gravedad [151] permanecen viables siempre y cuando la masa del gravitón sea pequeña y la materia se acople mínimamente a una de las métricas.

Probando la gravedad con la distancia luminosidad de OG

Otra prueba fundamental de la propagación de OG es su amplitud, que se escala con la inversa de la distancia desde la fuente hasta el observador. La RG predice que la distancia luminosidad de OG $d_L^{\rm gw}$ es igual a la distancia de luminosidad electromagnética $d_L^{\rm em}$, que sobre fondos cosmológicos es solo una integral de la longitud de Hubble (4.0.2). Las teorías modificadas de la gravedad modifican genéricamente esta predicción. La relación $d_L^{\rm gw}/d_L^{\rm em}$ se convierte así en un poderoso discriminador de cualquier término de fricción adicional a lo largo de la propagación.

Hemos determinado cómo $d_L^{\rm gw}$ diferiría de $d_L^{\rm em}$ en teorías genéricas de la gravedad cuando hay una modificación de la absorción del medio cósmico y una velocidad de OG anómala. Nuestro resultado principal se encuentra en la ecuación (4.1.17). Cuando $c_g = c$, la proporción $d_L^{\rm gw}/d_L^{\rm em}$ aún no es uno. Esto sugiere una vía prometedora para probar los modelos de EO que sobrevivieron a GW170817.

La distancia de luminosidad de OG se puede probar con OG, usando conteos de eventos, o con eventos multi-mensajeros, usando sirenas estándar u oscuras. Hemos contribuido a pronosticar las perspectivas de probar un $d_L^{\rm gw}$ modificada con sirenas estándar en LISA. Estas fuentes son particularmente interesantes porque se espera que se detecten hasta redshifts $z \sim 5$. Por lo tanto, LISA tiene un poder significativamente mayor que los detectores actuales en la Tierra (ver Fig. 4.1). Usando estimaciones actualizadas para la configuración de LISA y su curva de sensibilidad, encontramos que $d_L^{\rm gw}/d_L^{\rm em}$ podrían al nivel de un punto porcentual ($1 - 4\,\%$ dependiendo de los supuestos en el catálogo simulado de AN binarios masivos).

Oscilaciones de OG como medidor de campos cosmológicos adicionales

Más allá del paradigma cosmológico estándar, los modelos de energía oscura típicamente introducen campos cosmológicos adicionales. Estos modos adicionales también se pueden probar con la propagación de las OG. Sobre fondos cosmológicos homogéneos e isótropos, las perturbaciones tensor solo se acoplan a otros modos tensor a nivel lineal. Un ejemplo paradigmático de una teoría con dos tipos de perturbaciones tensoriales es bigravity. No obstante, estos modos tensor adicionales podrían ser la configuración efectiva de otros tipos de campos, por ejemplo, campos gauge. En analogía a las oscilaciones de neutrinos, si hay modos tensoriales interactuando mientras se propagan, podría haber oscilaciones de ondas gravitacionales.

Hemos desarrollado un método para estudiar las oscilaciones de OG de forma general. Dividimos las posibles interacciones en masa, fricción, velocidad y mezcla quiral. Todos ellas pueden producir oscilaciones en la amplitud de las OG, pero su relevancia relativa es diferente. En cada caso, identificamos la frecuencia de oscilación y comparamos la aproximación WKB y large-k. Curiosamente, encontramos que aunque las perturbaciones tensor acopladas a la materia se propagan a la velocidad de la luz, si el otro tensor tiene

una velocidad anómala y hay oscilaciones de OG, la velocidad de propagación efectiva de las OG puede diferir de la velocidad de la luz.

Las oscilaciones de OG dejan varias señales distintivas que hemos estudiado en detalle. Suponiendo que la producción de OG no se modifique, la mezcla de los modos tensoriales mientras se propagan se traduce en un déficit dependiente del tiempo de la OG original. Para una fuente dada, esta mezcla está fijada por la distancia y haría que la amplitud de la OG se redujera globalmente. Si la mezcla depende de la frecuencia de la onda, esto significaría una modulación dependiente del tiempo de la forma de onda. Estos dos efectos ocurren en modelos con fricción y mezcla en masa, respectivamente, como se ilustra en la Fig. 5.8. Pero si la señal es más tenue, esto se interpretaría como una fuente más distante. Por lo tanto, las oscilaciones de OG también modifican la distancia de luminosidad de OG de forma oscilatoria (ver Fig. 5.9). Este patrón distintivo se puede probar con varias sirenas estándar con diferentes desplazamientos al rojo. Tenga en cuenta también que en el caso de tener una conversión completa de un modo a otro, esto implicaría que no se detectarán OG en ciertos desplazamientos al rojo. Un efecto que es fácil de distinguir del modelado astrofísico. Además, si hay una velocidad anómala inducida, esto puede ser constreñido con un evento de multi-mensajero como GW170817. Finalmente, con una red de detectores se podría, en principio, detectar también una oscilación de OG dependiente de la polarización.

Hemos aplicado nuestro formalismo general para demostrar la capacidad de LISA para probar las oscilaciones de OG en modelos con interacciones de tensor adicionales, como bigravity. Calculamos las predicciones para bigravity utilizando una expansión de alta frecuencia, enfocándonos en el régimen de alta masa $m_g \gg H_0$. En esta región del espacio de parámetros, las oscilaciones de OG ocurren en el rango de mHz, pero la teoría no es capaz de explicar la aceleración cósmica. Las sirenas estándar a distancias cosmológicas, como será observado por LISA, tienen el potencial de restringir el rango de masa $m_g \gtrsim 10^{-25}$ eV para la mayoría de los ángulos de mezcla, 3 órdenes de magnitud mejores que las detecciones actuales de LIGO-Virgo. Esperamos que estas restricciones sean similares en otras teorías, por ejemplo, con una mezcla por fricción. Además, también anticipamos que la inclusión de efectos dependientes de la frecuencia en la forma de onda mejorará estos límites.

Agujeros negros primordiales como fósiles del universo primitivo

Yendo atrás en el tiempo, otro pilar del modelo cosmológico estándar son las perturbaciones de densidad primordiales que se convierten en semillas de la estructura del universo. Aún así, nuestro conocimiento de éstas es bastante limitado. En términos del espectro de potencia primordial (PS), solo las escalas más grandes se han visto limitadas hasta ahora, principalmente por el Fondo Cósmico de Microondas (FCM). Por lo tanto, las observaciones permiten grandes fluctuaciones a pequeña escala generadas en el universo temprano. Si fuera lo suficientemente grande (posiblemente del orden 10^{-1} en comparación con los 10^{-5} del FCM), tales fluctuaciones habrían colapsado al reingresar el horizonte para formar agujeros negros primordiales (ANP). Por lo tanto, si los

ANP existieran, constituirían una ventana nueva y directa a los períodos no probados del universo primitivo. Por el contrario, si no se encuentran, se pueden usar para colocar límites en diferentes escenarios de alta energía. Como bonus, los límites de la población de ANP también nos enseña cuánto pueden contribuir a la MO. Entre otros métodos, los ANP pueden estar limitados por los OG que emiten cuando se fusionan las binarias. Hemos estudiado la formación de ANP en modelos de inflación de un solo campo en el capítulo 6 y propusimos un escenario basado en critical Higgs inflation en el capítulo 7. También hemos determinado la relevancia de la difusión cuántica para la formación de ANPs desarrollando un método iterativo y calculando la cola de la distribución de perturbación de curvatura en los capítulos 8 y 9 respectivamente.

Formación de ANP la inflación de un solo campo

Una forma de producir ANPs es aumentar el PS primordial de perturbaciones de curvatura ζ. En la inflación de un solo campo esto puede lograrse teniendo una segunda meseta en el potencial. De esta manera, la inflación se ralentiza al cruzar el punto de inflexión y las perturbaciones de curvatura crecen ($\zeta = \delta\phi/\sqrt{2\epsilon}$ en el gauge de curvatura uniforme). Sin embargo, esta rápida desaceleración puede producir una violación de SR debido a $|\epsilon_2| > 1$.

Hemos reanalizado el aumento del espectro de potencia en modelos inflacionarios con un punto de inflexión. Más allá de SR, el PS podría crecer a escalas sub- o súper-horizonte. A escalas sub-horizonte, el PS crece cuando el campo se ralentiza desde $\mathcal{P}_\zeta \sim V(\phi)/\epsilon + \mathcal{O}(\epsilon^2)$. En las escalas de súper-horizonte, hay un modo constante y un modo evolutivo, que se amplifica o congela dependiendo del signo de la fricción. Cada vez que $2 - \epsilon + \epsilon_2 < 0$, el modo evolutivo crece exponencialmente, lo que solo se puede lograr si hay una violación de SR. Por lo tanto, durante este período fuera de SR, el PS también aumenta. Ambos mecanismos sirven para ampliar la producción de ANPs. Dependiendo del tiempo en que la inflación entre en la meseta y de cuánto tiempo la atraviese, el espectro de ANPs alcanzará su punto máximo en diferentes masas con un ancho diferente.

Un modelo motivado por la física de partículas: critical Higgs Inflation

Hemos propuesto un modelo motivado por la física de partículas en el que los ANP se forman desde un punto de inflexión en el potencial. En este escenario de critical Higgs inflation, el Higgs del modelo estándar acoplado no mínimamente a la gravedad actúa como el inflatón en el universo primitivo. Teniendo en cuenta la ecuación de grupo de renormalización (EGR) que se aplica tanto con el auto-acoplo del Higgs λ como el acoplamiento no mínimo con la gravedad ξ, encontramos regiones de espacio de parámetros permitidas por el Modelo estándar para el que el potencial de Higgs adquiere una segunda meseta a escalas más pequeñas, alrededor del punto crítico en el flujo de EGR $\lambda(\mu) \simeq \beta_\lambda(\mu) = 0$. Esta meseta proporciona una evolución de ultra SR del Higgs, que induce un pico alto y ancho en el espectro de potencia de curvatura.

Hemos explorado los observables del universo temprano y tardío de este modelo. Con respecto al FCM, encontramos regiones en el espacio de parámetros compatible

con Planck 2018 y mostrando una gran amplificación de las perturbaciones a escalas pequeñas. Sin embargo, los picos más grandes y más amplios tienden a tener un acuerdo peor, ya que el índice espectral se vuelve más pequeño. Curiosamente, este modelo predice una tensor-to-scalar ratio dentro del alcance de los experimentos actuales y de próxima generación, $r \sim 0{,}04$. En términos de la población de ANP, este escenario predice una distribución lognormal aproximada de masas. Su evolución depende de la fusión y el acrecentamiento, que está sujeto a la distribución inicial y la interacción con material bariónico. Bajo el supuesto de una distribución inicial altamente agrupada, encontramos que ANP podría constituir una fracción de la materia oscura con masas que van desde 0.01 a $100 M_{\odot}$.

Critical Higgs inflation y escenarios de ANP similares se pueden probar con OG de varias maneras. Una forma sería detectar un AN con una masa menor que la del sol, ya que no se conoce ningún mecanismo astrofísico que pueda formarlos. En este tipo de modelo, $0{,}1-1\,\%$ de las binarias de APNP tendrían un AN con una masa por debajo del límite de Chandrasekhar. Además, la acumulación de detecciones de fusión de binarias de AN puede restringir la distribución de masa. Probablemente, más prometedoras son las mediciones del spin efectivo, que los modelos de ANP generalmente predicen que son cercanos a cero, o las tasas de fusión a alto redshift, en las cuales los AN astrofísicos no tendrían tiempo para formarse. Finalmente, los escenarios de ANP como critical Higgs inflation producen fondos estocásticos tanto en la formación, por perturbaciones de segundo orden, como en la fusión, de las binarias no resueltas.

Implicaciones importantes de la difusión cuántica

Los ANP se generan a partir de grandes perturbaciones de curvatura. Sin embargo, las técnicas perturbativas aplicadas a la inflación no tienen en cuenta las contribuciones no lineales que pueden ser relevantes cuando $\zeta \sim 1$. Por esta razón, es más apropiado estudiar ANP utilizando una teoría de campo efectiva para las perturbaciones súper-horizonte en el que los modos de longitud de onda corta se comportan como un ruido estocástico clásico. Esto no es más que una inflación estocástica.

Hemos investigado la producción de ANP en modelos de inflación de un solo campo con un punto de inflexión, teniendo en cuenta la reacción de las fluctuaciones cuánticas. Dado que la producción de ANP está relacionada con una desviación de SR, hemos considerado el sistema de ecuaciones de Langevin asociado tanto al campo $\bar{\phi}$ como al momento $\bar{\pi}_{\phi}$ en el formalismo de Hamilton-Jacobi de la inflación estocástica. Para encontrar la función de densidad de probabilidad multivariable asociada a este problema estocástico, hemos resuelto el sistema de ecuaciones diferenciales de primer orden para los momentos estadísticos $\langle \delta\phi^m \delta\pi^n \rangle$. Con estos momentos, se puede construir la función característica $\chi(u_{\phi}, u_{\pi})$ que se puede usar para calcular la función de densidad de probabilidad (PDF). La metodología que hemos desarrollado se puede aplicar para estudiar los efectos de la difusión cuántica más allá de SR en general. Hemos particularizado su uso para encontrar la probabilidad de que una fluctuación dada esté por encima de un cierto umbral $P(\zeta > \zeta_c)$, que es la pregunta relevante para saber si se forman ANPs.

Hemos analizado las implicaciones de la difusión cuántica más allá de SR para la

producción de ANP. Los modelos inflacionarios con un punto de inflexión inducen un crecimiento exponencial de los modos de curvatura siempre que $3 - \epsilon + \epsilon_2 < 0$. Encontramos que este aumento de las fluctuaciones cuánticas también aumenta el ruido estocástico. A su vez, esto produce un pico en el espectro de potencia sobre el espectro de ζ_k sin difusión. Tal pico de difusión tiene un efecto claro que amplifica la producción de ANP. Además, observamos que este gran ruido estocástico también induce importantes contribuciones no Gaussianas. Estas contribuciones no Gaussianas podrían aumentar o disminuir la producción de ANP. Para nuestro potencial, vemos que la contribución dominante está dada por el cuarto momento. Introduce una contribución positiva que aumenta el peso de las colas de la PDF, lo que lleva a una probabilidad significativamente mayor de formar ANPs. Sumando ambas contribuciones, encontramos que para nuestro caso de estudio, la difusión cuántica aumenta sustancialmente la formación de ANPs.

Nuestros resultados muestran que la difusión cuántica más allá de SR puede tener un papel importante en la determinación de la abundancia de ANP para modelos de campo único con un punto de inflexión. Estos efectos se pueden parametrizar con la dinámica de los parámetros de flujo de Hubble $\epsilon_n(N)$. Esperamos que nuestras conclusiones puedan extenderse a diferentes potenciales estudiados en la literatura que comparten un comportamiento similar de ϵ_n. Creemos que las predicciones para el espectro de ANP en esta clase de modelos deben revisarse en el futuro para la difusión cuántica.

El papel de las colas de la distribución

La abundancia de ANP viene dada por la probabilidad de que una perturbación de curvatura esté por encima de un cierto umbral. Por lo tanto, la producción de ANP está controlada por la cola de la PDF $P(\zeta)$. Hemos visto que las correcciones no Gaussianas inducidas por el ruido estocástico pueden ser relevantes en el punto de inflexión. Entonces sería relevante tener técnicas no perturbativas para caracterizar completamente la cola sin necesitar una expansión en los momentos estadísticos.

Hemos desarrollado nuevas técnicas para calcular la cola de la PDF en el contexto de la inflación de SR. Demostramos que la caída exponencial de la cola está controlada por los polos en la función característica. Estos polos pueden computarse numéricamente o analíticamente en ciertos casos, como el potencial lineal y casi constante que estudiamos. Alternativamente, uno podría obtener la cola de la PDF a partir de los valores propios de un problema equivalente equivalente derivado de la ecuación de Fokker-Planck. Encontramos que los polos y los valores propios dependen de los parámetros del potencial y de las condiciones de contorno. Es decir, del rango de posibles valores de campo iniciales considerados de $\phi_{\rm uv}$ a $\phi_{\rm end}$.

También investigamos potenciales dominados por vacío más realistas con una pendiente polinómica. Allí, estamos interesados en el régimen en el que la cola de la PDF no es sensible a las condiciones de contorno en altas energías. Encontramos que la cola de la PDF derivada de la evolución estocástica puede ser muy bien aproximada por un potencial plano en la región donde domina la difusión cuántica, determinada por $v^2 v''/v'^2 \gtrsim 1$. Comparamos los resultados numéricos de los polos con esta aproximación

de pozo cuántico obteniendo un buen acuerdo. Esto implica que la cola de la PDF es una propiedad universal para estos potenciales.

Concluimos que la cola de la PDF es significativamente más relevante que la estimación Gaussiana habitual. Esto se debe a que la difusión cuántica induce un decaimiento exponencial lineal a la curvatura $\sim e^{-\Lambda_n \zeta}$ en lugar de $\sim e^{-\zeta^2/2\sigma^2}$. Por supuesto, esto también implica que para un umbral dado ζ_c, la cantidad de ANP formado es muchos órdenes de magnitud mayor que la predicción Gaussiana. Esto tiene implicaciones importantes para la construcción de modelos inflacionarios y la fenomenología de ANP.

Part V

Appendices

CHAPTER

A

A mathematical detour: modified gravity in the language of differential forms

Dark energy is intimately related to our understanding of gravity on large scales. We have seen in part II that testing the propagation of GWs is a very powerful tool to unveil its properties. From the theoretical side, it is essential to explore the most general healthy gravity theories and their possible phenomenological fingerprints, determining as well how the physical predictions of different theories beyond GR relate to each other.

Motivated by the necessity of finding unifying frameworks, we developed in [1] a novel formalism to construct gravity theories based on the language of differential forms. By the antisymmetric properties of these mathematical objects, they are perfectly suited to build theories with second order equations of motion, for example Massive Gravity or Horndeski theory. Moreover, we found that different gravity theories can be naturally connected by field redefinition in this language [3]. These works set a new basis to explore alternatives to GR beyond previous analysis, being especially useful for geometrical approaches to gravity. For instance, when gravity emerges from a higher-dimensional geometry (i.e. Kaluza-Klein reduction) or when torsion is non-zero (when fermions are present). In addition, differential forms can be very efficiently implemented in computer codes for tensor algebra such as `xAct` for `MATHEMATICA`, showing also the practical advantage of this new framework. In the following we summarize the main aspects of the differential form formalism.

A.1 Towards the most general scalar-tensor theories of gravity

Scalar-tensor theories are generally described by an action functional S, which corresponds to the integral of the Lagrangian $\mathcal{L}$ over the curved space-time. We are going to exploit the fact that, mathematically, integration is an operation defined in terms of the space of differential forms $\Omega^q(\mathcal{M})$, where q is the order of the q-form and the dimension of the base manifold $\mathcal{M}$. Since the action is defined as an integral over a D-dimensional curved space-time manifold, the Lagrangian must be a D-form, i.e.

$$S = \int_{\mathcal{M}} \mathcal{L}. \tag{A.1.1}$$

Crucially, a D-form is characterized for being proportional to the volume element $\eta = \sqrt{-g}dx^1 \wedge \cdots \wedge dx^D$, leading to a direct connection with the usual component notation. Furthermore, due to the fact that $\Omega^q(\mathcal{M})$ is constructed as the space of totally antisymmetric $(0,q)$-tensors, if we construct our D-form Lagrangians with exterior products of differential forms, the set of possibilities will be *finite*.

In order to determine a general basis for scalar-tensor Lagrangians, we must first identify the appropriate building blocks written in differential form language. From the tensorial side, we have the usual geometrical quantities characterizing a manifold. In particular, we will work with differentiable manifolds with an associated metric $\mathfrak{g}$ and 1-form connection $\omega^a{}_b$. Also, we will fix the metric to have a Lorentzian signature. Moreover, we will focus on manifolds with a vanishing torsion $T^a = 0$ and a metric-compatible connection $\omega_{ab} = -\omega_{ba}$, i.e. *pseudo-Riemannian manifolds*. In such a case, the connection is uniquely determined by the non-coordinate basis elements θ^a, which can be related to the curved space-time metric via the flat Minkowski metric η_{ab}, i.e. $\mathfrak{g} = \eta_{ab}\theta^a \otimes \theta^b$. Introducing an exterior covariant derivative $\mathcal{D}$ constructed from $\omega^a{}_b$, the geometry of the manifold is encoded in the 2-form curvature, defined as $\mathcal{R}^a{}_b = \mathcal{D}\omega^a{}_b$. This will be our building block characterizing the tensorial part of the action. In components, it reads

$$\mathcal{R}^a{}_b = \frac{1}{2}R^a{}_{bcd}\theta^c \wedge \theta^d, \tag{A.1.2}$$

where $R^a{}_{bcd}$ is the corresponding Riemann tensor. One should notice that, throughout the text, we will use latin indices to denote non-coordinate components and greek indices for coordinate ones. Both basis are linked with the vielbein $e^a{}_\mu$ by $\theta^a = e^a{}_\mu dx^\mu$. Moreover, the 1-form connection ω and the Levi-Civita connection Γ are related by the vielbein postulate $\nabla_\mu e^a{}_\nu = 0$. In this language, Bianchi's second identity simply implies that $\mathcal{D}\mathcal{R}^a{}_b = 0$. For a review on differential geometry, we recommend [493].

Subsequently, we must encounter possible q-forms describing the scalar field and its derivatives. The scalar field ϕ itself defines a 0-form. Its partial derivative is also a well-defined 1-form, corresponding to the exterior derivative of the scalar field $d\phi = \nabla_\mu\phi dx^\mu$. However, it is not trivial to introduce the second covariant derivative of the scalar field $\nabla_\mu\nabla_\nu\phi$ because it is a symmetric $(0,2)$-tensor. Consequently, we must find an appropriate antisymmetric tensor which encodes the information from the second derivatives. Since the tensor is symmetric, we cannot apply directly an antisymmetric operator, i.e. $\nabla_{[\mu}\nabla_{\nu]}\phi = 0$. If we apply an antisymmetric operator to only one of the indices, in order to finally obtain a q-form, we will end up with a D-form, which is a trivial case since it is already proportional to the volume element. Additionally, using Poincare lemma, the exterior derivative of the gradient field vanishes, i.e. $dd\phi = 0$. Moreover, by definition, the wedge product of $d\phi$ with itself is also zero, i.e. $d\phi \wedge d\phi = 0$. This means that using this 1-form we could never construct the kinetic term, because it contains two first derivatives. Clearly, we need more adequate definitions of the q-forms representing the first and second derivatives of the scalar field. In the following, we propose a minimal setup, in which derivatives of the field appear in the lowest possible order while fulfilling our requirements. This leads to two derivatives of the scalar in each

element of the basis.

Let us define two vector-valued 1-forms that encode the first and second covariant derivatives of ϕ

$$\Psi^a \equiv \nabla^a\phi\nabla_b\phi\ \theta^b, \tag{A.1.3}$$

$$\Phi^a \equiv \nabla^a\nabla_b\phi\ \theta^b. \tag{A.1.4}$$

Then, we will construct the most general scalar-tensor theory obeying the following:

- It is described by an action principle in which the Lagrangian is a D-form invariant under Local Lorentz Transformations (LLT) defined in a pseudo-Riemannian manifold.
- The Lagrangian is built up out of exterior products of the vielbein θ^a, the 2-form curvature $\mathcal{R}^{ab}$, first derivatives of the scalar field Ψ^a and second derivatives of the scalar field Φ^a.

As a consequence, in order to have a Lagrangian invariant under LLT, there cannot be free indices. Thus, they must be contracted with the tangent space metric η_{ab} and the totally antisymmetric symbol $\epsilon_{a_1\cdots a_D}$, which are invariant objects[1]. Moreover, the fact that we restrict to pseudo-Riemannian manifolds, i.e. manifolds with a metric-compatible connection and a vanishing torsion, implies that all the tensorial dynamics is contained in the 2-form curvature (A.1.2). With these two conditions, we can define a basis of Lagrangian given by

$$\mathcal{L}_{(lmn)} = \bigwedge_{i=1}^{l}\mathcal{R}^{a_ib_i}\wedge\bigwedge_{j=1}^{m}\Phi^{c_j}\wedge\bigwedge_{k=1}^{n}\Psi^{d_k}\wedge\theta^\star{}_{a_1b_1\cdots a_lb_lc_1\cdots c_md_1\cdots d_n}, \tag{A.1.5}$$

where $\bigwedge$ is an abbreviation for a set of consecutive wedge products and $l, m, n \in \mathbb{N}$. In this notation, if any of the subindices of the Lagrangian are zero, the corresponding terms in the r.h.s do not appear. Here, $\theta^\star_{a_1\cdots a_k}$ is the Hodge dual basis and it is defined as

$$\theta^\star{}_{a_1\cdots a_k} = \frac{1}{(D-k)!}\epsilon_{a_1\cdots a_ka_{k+1}\cdots a_D}\theta^{a_{k+1}}\wedge\cdots\wedge\theta^{a_D}. \tag{A.1.6}$$

One should notice that the previous result $d\phi\wedge d\phi = 0$ appears in this notation making $\mathcal{L}_{(lmn)}$ vanish for $n > 1$. Additionally, it must hold that $2l + m + n \leqslant D$ due to the antisymmetry by the Hodge dual basis. This will be very important because it means that for a given dimension D our basis of Lagrangians will be finite. Interestingly, if we do not include the scalar field, setting $m = n = 0$, these Lagrangians correspond to Lovelock's theory [38] written in differential forms. Therefore, this basis of Lagrangians could be seen as its *scalar-tensor extension*. Finally, it is important to remark that there are three additional Lagrangians that fulfill our premises but are not included in

[1]We will choose the convention $\eta_{ab} = \mathrm{diag}(-1, 1, 1, 1)$ for the metric signatures and $\epsilon_{0123} = +1$ for the antisymmetrizations.

our basis (A.1.5). They correspond to Lagrangians in which the indices of the building blocks are contracted among them, e.g. $\mathcal{R}^{ab} \wedge \Phi_a \wedge \Psi_b$. However, they do not lead to second order equations of motion. Thus, we discard them from the beginning.

In the scalar-tensor theories represented by the basis (A.1.5), the action will be the sum over all possible Lagrangians with different l, m and n integrated over the space-time manifold, i.e.

$$S = \sum_{l,m,n}^{p\leqslant D} \int_{\mathcal{M}} \alpha_{lmn} \mathcal{L}_{(lmn)}, \tag{A.1.7}$$

where $p \equiv 2l + m + n$ and $n \leqslant 1$. In this context, the coefficients α_{lmn} represent 0-forms, which, in general,[2] can be functions of the scalar field and its derivatives $\alpha_{lmn} = \alpha_{lmn}(\phi, X, [\Phi], \cdots)$, where we are using the notation for which a square bracket represents the contraction of two free indices, e.g. $[t_{\mu\nu}] \equiv t^{\mu}{}_{\mu}$, and an angle bracket the contraction with partial derivatives of the scalar field, e.g. $\langle t_{\mu\nu} \rangle \equiv \phi^{,\mu} t_{\mu\nu} \phi^{,\nu}$. Also, partial derivatives are shortened by a comma, $\partial_\mu \phi = \phi_{,\mu}$, and covariant derivatives are shortened by a semicolon, $\nabla_\mu \nabla_\nu \phi = \phi_{;\mu\nu}$. Lastly, we write the contractions of second derivatives as $\Phi^n{}_{\mu\nu} = \phi_{;\mu\alpha_1} \phi^{;\alpha_1}{}_{;\alpha_2} \cdots \phi^{;\alpha_{n-1}}{}_{;\nu}$ and define $-2X \equiv \phi^{,\mu}\phi_{,\mu}$.

In $4D$, we have 15 possible Lagrangians in our basis. In order to translate them into the usual component notation, we only need to apply the definition of the wedge product and the Hodge dual basis. Recalling that $\eta = \theta^1 \wedge \cdots \wedge \theta^D$ is the volume element, we find the following Lagrangians:

(i) $p = 0$

$$\mathcal{L}_{(000)} = \theta^\star = \eta, \tag{A.1.8}$$

(ii) $p = 1$

$$\mathcal{L}_{(010)} = \Phi^a \wedge \theta^\star{}_a = [\Phi] \cdot \eta, \tag{A.1.9}$$

$$\mathcal{L}_{(001)} = \Psi^a \wedge \theta^\star{}_a = -2X \cdot \eta, \tag{A.1.10}$$

(iii) $p = 2$

$$\mathcal{L}_{(100)} = \mathcal{R}^{ab} \wedge \theta^\star{}_{ab} = R \cdot \eta, \tag{A.1.11}$$

$$\mathcal{L}_{(020)} = \Phi^a \wedge \Phi^b \wedge \theta^\star{}_{ab} = ([\Phi]^2 - [\Phi^2])\eta, \tag{A.1.12}$$

$$\mathcal{L}_{(011)} = \Phi^a \wedge \Psi^b \wedge \theta^\star{}_{ab} = -(\langle \Phi \rangle + 2X[\Phi])\eta, \tag{A.1.13}$$

(iv) $p = 3$

$$\mathcal{L}_{(110)} = \mathcal{R}^{ab} \wedge \Phi^c \wedge \theta^\star{}_{abc} = -2G^{ab}\Phi_{ab}\eta, \tag{A.1.14}$$

$$\mathcal{L}_{(030)} = \Phi^a \wedge \Phi^b \wedge \Phi^c \wedge \theta^\star{}_{abc} = ([\Phi]^3 - 3[\Phi][\Phi^2] + 2[\Phi^3])\eta, \tag{A.1.15}$$

$$\mathcal{L}_{(101)} = \mathcal{R}^{ab} \wedge \Psi^c \wedge \theta^\star{}_{abc} = -2\langle G \rangle \eta, \tag{A.1.16}$$

$$\mathcal{L}_{(021)} = \Phi^a \wedge \Phi^b \wedge \Psi^c \wedge \theta^\star{}_{abc} = 2(\langle \Phi^2 \rangle - \langle \Phi \rangle [\Phi] - X([\Phi]^2 - [\Phi^2]))\eta, \tag{A.1.17}$$

[2]Here, it will be important that the coefficient is a 0-form and that we are constructing the geometrical quantities out of the 2-form curvature $\mathcal{R}^{ab}$. Consequently, we will not consider any dependence in curvature scalars in α_{lmn}, e.g. R^2 or $R_{ab}R^{ab}$.

(v) $p = 4$

$$\mathcal{L}_{(200)} = \mathcal{R}^{ab} \wedge \mathcal{R}^{cd} \wedge \theta^{\star}{}_{abcd} = (R_{abcd}R^{abcd} - 4R_{ef}R^{ef} + R^2)\eta, \tag{A.1.18}$$

$$\begin{aligned}\mathcal{L}_{(120)} &= \mathcal{R}^{ab} \wedge \Phi^c \wedge \Phi^d \wedge \theta^{\star}{}_{abcd} \\ &= (R([\Phi]^2 - [\Phi^2]) - 4R^{ab}([\Phi]\Phi_{ab} - \Phi^2_{ab}) + 2R^{abcd}\Phi_{ac}\Phi_{bd})\eta,\end{aligned} \tag{A.1.19}$$

$$\begin{aligned}\mathcal{L}_{(040)} &= \Phi^a \wedge \Phi^b \wedge \Phi^c \wedge \Phi^d \wedge \theta^{\star}{}_{abcd} \\ &= ([\Phi]^4 - 6[\Phi]^2[\Phi^2] + 3[\Phi^2]^2 + 8[\Phi][\Phi^3] - 6[\Phi^4])\eta,\end{aligned} \tag{A.1.20}$$

$$\begin{aligned}\mathcal{L}_{(111)} &= \mathcal{R}^{ab} \wedge \Phi^c \wedge \Psi^d \wedge \theta^{\star}{}_{abcd} \\ &= \left(4\left(\langle R_{ab}\Phi^{bc}\rangle + X[R\Phi]\right) - R\left(\langle\Phi\rangle + 2X[\Phi]\right) + 2\left(\langle R_{abcd}\Phi^{bd}\rangle - \langle R\rangle[\Phi]\right)\right)\eta,\end{aligned} \tag{A.1.21}$$

$$\begin{aligned}\mathcal{L}_{(031)} &= \Phi^a \wedge \Phi^b \wedge \Phi^c \wedge \Psi^d \wedge \theta^{\star}{}_{abcd} \\ &= (6(\langle\Phi^2\rangle[\Phi] - \langle\Phi^3\rangle) - 3\langle\Phi\rangle([\Phi]^2 - [\Phi^2]) - 2X([\Phi]^3 - 3[\Phi][\Phi^2] + 2[\Phi^3]))\eta,\end{aligned} \tag{A.1.22}$$

where R is the Ricci scalar, R_{ab} is the Ricci tensor and G_{ab} is the Einstein tensor, given by $G_{ab} = R_{ab} - \frac{1}{2}g_{ab}R$. As a consequence of the above expressions, we can easily relate our results with the current literature. For instance, the modern version of Horndeski's Theory [65] is a linear combination of (A.1.8), (A.1.9), (A.1.11-A.1.12) and (A.1.14-A.1.15), and the class of viable theories Beyond Horndeski known as Generalized Generalized Galileons (G^3) [100] are simply (A.1.17) and (A.1.22). In addition, terms such as (A.1.16) and (A.1.21) appear when doing a Kaluza-Klein compactification of higher dimensional Lovelock's densities [494] and correspond respectively to "John" and "Paul" Lagrangians of the Fab Four theory [93]. Furthermore, when we are in flat space, Galileon theory [44] is built up with (A.1.9), (A.1.10), (A.1.13), (A.1.17) and (A.1.22).

In [1], we showed that there is a well-established interconnection between all these Lagrangians. In fact, not all of them are independent and only certain linear combinations give rise to second order equations of motion. A summary of these relations among Lagrangians can be found in Fig. A.1. Here we skip the details on how to compute the equations of motions and determine the completeness of the basis of Lagrangians in this formalism. We find two sets of Lagrangians $\mathcal{L}_i^H$ and $\mathcal{L}_i^{NH}$ with second order EoM. We summarize the first set of Lagrangians in

$$\mathcal{L}_2^H[G_2] = G_2\mathcal{L}_{(000)}, \tag{A.1.23}$$

$$\mathcal{L}_3^H[G_3] = G_3\mathcal{L}_{(010)}, \tag{A.1.24}$$

$$\mathcal{L}_4^H[G_4] = G_4\mathcal{L}_{(100)} + G_{4,X}\mathcal{L}_{(020)}, \tag{A.1.25}$$

$$\mathcal{L}_5^H[G_5] = G_5\mathcal{L}_{(110)} + \frac{1}{3}G_{5,X}\mathcal{L}_{(030)}, \tag{A.1.26}$$

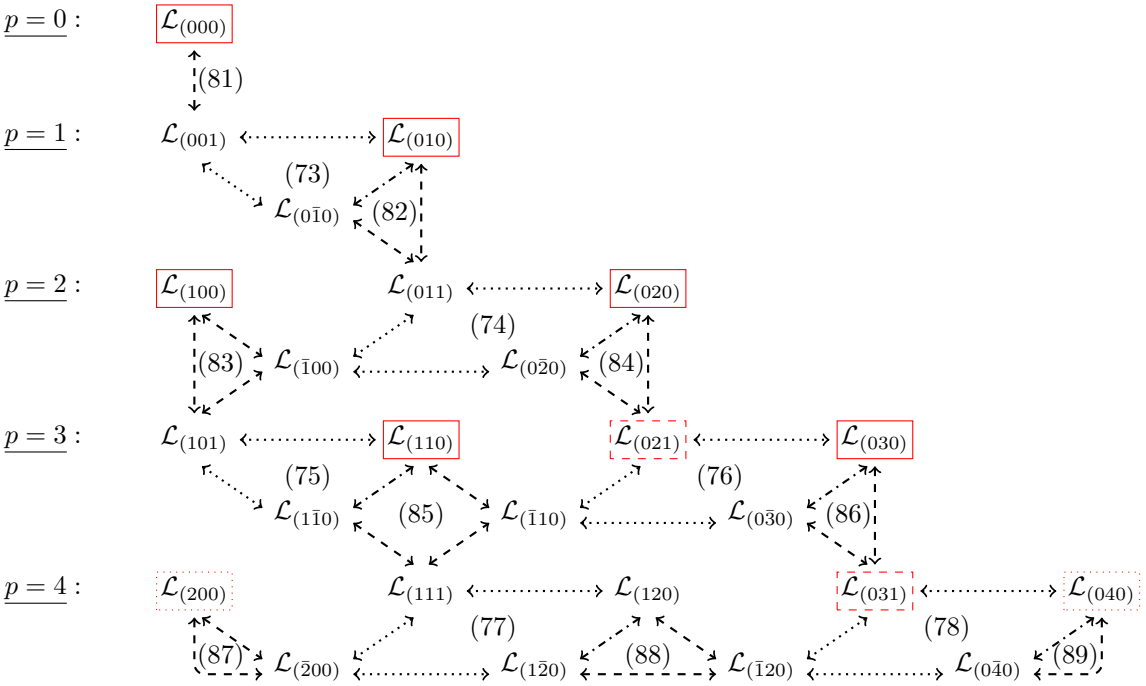

Figure A.1. Summary of the interconnections between different Lagrangians $\mathcal{L}_{(lmn)}$, $\mathcal{L}_{(\bar{l}mn)}$ and $\mathcal{L}_{(l\bar{m}n)}$. A close set of arrows indicates that the Lagrangians in the vertices are related by the identity referred in the interior, which can be either an exact form, presented with dotted arrows, or an algebraic antisymmetric identity, plotted with dashed arrows. Here, a dash-dotted arrow indicates that two Lagrangians are related by both types of identities. In total, there are 10 independent Lagrangians. In the figure, we show a possible choice, framing each term in a rectangle, corresponding to Horndeski theory (red rectangles), Beyond Horndeski's G^3 (red dashed rectangles) and $\mathcal{L}_{(200)}$ and $\mathcal{L}_{(040)}$ (red dotted rectangles). Finally, we emphasize the structure by levels indicating in the left the number of building blocks $p \equiv 2l + m + n$. Find more details in [1].

which is nothing but the differential form version of Horndeski's theory. We can englobe the whole set in $\mathcal{L}^H = \sum_{i=2}^{5} \mathcal{L}_i^H$. The second set we found was

$$\mathcal{L}_2^{NH}[E_2] = E_2 \mathcal{L}_{(001)}, \tag{A.1.27}$$

$$\mathcal{L}_3^{NH}[E_3] = E_3 \mathcal{L}_{(011)}, \tag{A.1.28}$$

$$\mathcal{L}_4^{NH}[E_4] = E_4 \mathcal{L}_{(101)} + E_{4,X} \mathcal{L}_{(021)}, \tag{A.1.29}$$

$$\mathcal{L}_5^{NH}[E_5] = E_5 \mathcal{L}_{(111)} + \frac{1}{3} E_{5,X} \mathcal{L}_{(031)}, \tag{A.1.30}$$

$$\mathcal{L}_6^{NH}[E_6] = E_6 \mathcal{L}_{(200)} + 2E_{6,X} \mathcal{L}_{(120)} + \frac{1}{3} E_{6,XX} \mathcal{L}_{(040)}. \tag{A.1.31}$$

Consequently, the aspect that we need to address is if Eqs. (A.1.27-A.1.31) contain any dynamics beyond Eqs. (A.1.23-A.1.26).

We conclude that there is a total of ten independent Lagrangians, which can be chosen to be the six of Horndeski, plus Beyond Horndeski, i.e. $\mathcal{L}_{(021)}$ and $\mathcal{L}_{(031)}$, plus $\mathcal{L}_{(040)}$ and the Gauss-Bonnet $\mathcal{L}_{(200)}$. From them, there are only four independent combinations giving rise to second order EoM These four independent Lagrangians can be chosen to be the ones of Horndeski, i.e. (A.1.23-A.1.26). We realize that with this procedure we are not able to conclude anything whether $\mathcal{L}_{(021)}$ and $\mathcal{L}_{(031)}$ are well behaved by themselves, as they do in Beyond Horndeski theories (G^3) [100]. This would require a Hamiltonian analysis. However, this result tells us that the higher derivative structure of G^3 model, i.e. $\mathcal{L}_{(021)}$ and $\mathcal{L}_{(031)}$, is precisely the same as the one of $\mathcal{L}_{(101)}$ and $\mathcal{L}_{(111)}$ respectively. This seems to indicate that those terms might also be ghost free.

In the field of general scalar-tensor theories, the fundamental analysis was made by Horndeski [65], who found the most general second order scalar-tensor Euler-Lagrange equations in four dimensions. In practice, Horndeski's theorem was first proven at the level of the equations of motion, imposing a relation between the divergence of the metric and the scalar field equations arising from diffeomorphism invariance. He then classified all the possible terms compatible with this requirement and proceeded by finding an action that produced them in the equations of motion. In this sense, our work has followed the opposite direction. We have started by looking for the most general action satisfying invariance under Local Lorentz Transformations in a pseudo-Riemannian manifold and constructed with a fixed set of building blocks; the vielbein θ^a, the curvature 2-form $\mathcal{R}^{ab}$, the 1-form Ψ^a encoding first derivatives of the scalar field and the 1-form Φ^a containing second derivatives linearly. Then, we have looked for the combinations which give rise to second-order EoM It is important to remark that in this paper we have not proven Horndeski's theorem, since our basis of Lagrangians can be generalized to higher powers of the derivatives of the field. However, what we *have proven* is that Horndeski's theory corresponds to the most general second order 4-form Lagrangian invariant under LLT in a pseudo-Riemannian manifold and constructed with θ^a, $\mathcal{R}^{ab}$, Ψ^a and Φ^a. Consequently, Horndeski theorem guarantees that any non-linear extension of our basis will be either equivalent to it or characterized by higher derivatives EoM

It is interesting to note that the Lagrangian $\mathcal{L}_6^{NH}[E_6(\phi, X)]$, which we named *kinetic Gauss-Bonnet* and presented in (A.1.31), has not been previously studied in the literature. However, we have also shown in this work that its dynamics is already contained in the full Horndeski's theory. A particular case of this Lagrangian, when $E_6 = E_6(\phi)$, is the well-known scalar coupling to the Gauss-Bonnet term $f(\phi)GB$ [376]. In this respect, with the previous result, we have additionally proven *explicitly* that such a theory belongs to Horndeski, as it was claimed in Ref. [86]. Interestingly, when there is only kinetic dependence in the coefficient, i.e. $E_6 = E_6(X)$, the kinetic Gauss-Bonnet Lagrangian becomes identically an exact form.

A.2 Field redefinitions in theories beyond Einstein gravity

Gravity can be easily formulated in the tangent space. In this context, the usual diffeomorphism (Diff) invariance becomes an invariance under local Lorentz Transformations (LLT), resembling in a clear manner the similarities with gauge theories. Moreover, the geometry of the manifold, which was previously determined by the space-time components of the metric tensor $g_{\mu\nu}$, is now contained in the basis elements of the cotangent space θ^a. The two objects are directly connected through the definition of the metric tensor $\mathfrak{g} = g_{\mu\nu}dx^\mu \otimes dx^\nu = \eta_{ab}\theta^a \otimes \theta^b$, where η_{ab} is the Minkowski metric, given by $\eta_{ab} = \text{diag}(-1, 1, 1, 1)$. If one wants to transform the metric while maintaining its causal structure, we simply need to rescale the basis elements θ^a by $\tilde{\theta}^a = \tilde{C}(x)\theta^a$, which is nothing but a *conformal transformation*. In the case of JBD theories [72], one could choose the conformal factor to depend on the scalar field ϕ in order to eliminate the non-minimal coupling with the Ricci scalar. This transformation, in the usual component notation, takes the well-known form $\tilde{g}_{\mu\nu} = \tilde{C}^2 g_{\mu\nu}$.

However, when there are derivative couplings of the scalar field to the curvature, such as in Horndeski's theory [65], conformal transformations are not enough to erase the non-minimal couplings. Nonetheless, one could take advantage of the results of previous section, where it was shown that scalar-tensor theories can be naturally built using the language of differential forms. Apart from the curvature 2-form describing the geometry, one just needs two 1-forms, $\Psi^a \equiv \nabla^a\phi\nabla_b\phi\theta^b$ and $\Phi^a \equiv \nabla^a\nabla_b\phi\theta^b$, encoding respectively first and second derivatives of the scalar field. Thus, one could perform a redefinition of the vielbein θ^a that includes first derivatives of the scalar field by applying

$$\tilde{\theta}^a = \tilde{C}(\phi, X)\theta^a + \tilde{D}(\phi, X)\Psi^a, \tag{A.2.1}$$

where X is the scalar kinetic term $-2X = \nabla_\mu\phi\nabla^\mu\phi$. This kind of field redefinition is known as a *disformal transformation* [97]. In order for this transformation to be well behaved, it must have a non-vanishing determinant, so that it can be inverted. In the language of differential forms, the determinant of the transformation can be very easily computed via the volume element, which encodes it naturally with a square root. Recalling that it is given by $\eta \equiv \frac{1}{D!}\epsilon_{a_1\cdots a_D}\theta^{a_1} \wedge \cdots \wedge \theta^{a_D} = \sqrt{-g}dx^D$, we obtain that it transforms as

$$\tilde{\eta} = \frac{1}{D!}\epsilon_{a_1\cdots a_D}\tilde{\theta}^{a_1} \wedge \cdots \wedge \tilde{\theta}^{a_D} = \tilde{C}^{D-1}\left(\tilde{C} - 2X\tilde{D}\right)\eta, \tag{A.2.2}$$

where $\epsilon_{a_1\cdots a_D}$ is the totally antisymmetric symbol[3]. Therefore, to prevent the disformal volume element to become a complex number, the disformal coefficients must satisfy that $\tilde{C} > 0$ and $\tilde{C} > 2X\tilde{D}$, where the transformed determinant arises from $\tilde{\eta} = \sqrt{-\tilde{g}}\eta$.

These disformal transformations in the tangent space can be traced back to the conventional component notation. Starting from the definition of the disformal metric, we find that

$$\tilde{\mathfrak{g}} = \eta_{ab}\tilde{\theta}^a \otimes \tilde{\theta}^b = \tilde{g}_{\mu\nu}dx^\mu \otimes dx^\nu = \left(\tilde{C}^2 g_{\mu\nu} + 2\tilde{D}\left(\tilde{C} - X\tilde{D}\right)\nabla_\mu\phi\nabla_\nu\phi\right)dx^\mu \otimes dx^\nu. \tag{A.2.3}$$

[3]This relation can be trivially obtained using that $\Psi^a \wedge \Psi^b = 0$, which is a consequence of the antisymmetry of the exterior product.

Subsequently, one can recover the original formulation of Bekenstein [97] by appropriately redefining the disformal coefficients.

The benefits of applying disformal transformations in the tangent space are considerable. First of all, it naturally introduces the 1-form Ψ^a, which was one of the basic building blocks used in (A.1.5). Thus, it connects in an interesting and fundamental manner this new formulation for ST theories with disformal transformations. Secondly, the building blocks of the ST theories will transform in a very transparent way. From the transformed frame field $\tilde{\theta}^a$, the rest of geometrical quantities, the connection 1-form $\tilde{\omega}^{ab}$ and the 2-form curvature $\tilde{\mathcal{R}}^{ab}$, can be constructed. For the scalar field building blocks, the transformation also follows directly. Therefore, this method provides us with a way to compute the disformal building blocks $\tilde{\mathcal{R}}^{ab}$, $\tilde{\Psi}^a$ and $\tilde{\Phi}^a$ in terms of the original ones Ψ^a, Φ^a and $\mathcal{R}^{ab}$. This fact will simplify enormously the computations because any disformally transformed Lagrangian could be expressed as a linear combination of Lagrangians with the same building blocks. Moreover, it will become straightforward to elucidate how each new Lagrangian is generated through the specific dependence of the disformal coefficients on ϕ and X.

In order to study the disformal transformation of a gravity theory, we first need to investigate how the building blocks transform. For example, we can obtain the disformal connection 1-form from the torsionless and metricity conditions on the disformal connection, i.e. $\tilde{T}^a = 0$ and $\tilde{\omega}_{ab} = -\tilde{\omega}_{ba}$. We postulate that the disformal connection takes the form $\tilde{\omega}^a{}_b = \omega^a{}_b + X^a{}_b$, where $X^a{}_b$ must satisfy $X_{ab} = -X_{ba}$. Then, it can be determined from

$$\tilde{T}^a = \tilde{\mathcal{D}}\tilde{\theta}^a = d\tilde{\theta}^a + \tilde{\omega}^a{}_b \wedge \tilde{\theta}^b = \mathcal{D}\tilde{\theta}^a + X^a{}_b \wedge \tilde{\theta}^b \,, \tag{A.2.4}$$

where $\mathcal{D}$ represents an exterior, covariant derivative. After implementing the torsionless condition, the 1-form connection follows

$$\tilde{\omega}^{ab} = \omega^{ab} - (\Omega_\phi + \frac{1}{2}\langle\Phi\rangle\Omega_X\Upsilon)\cdot\theta^{[a}\nabla^{b]}\phi + \Omega_X\cdot\theta^{[a}\nabla^{b]}\nabla_c\phi\nabla^c\phi + \Upsilon\cdot\Phi^{[a}\nabla^{b]}\phi + \Upsilon_X\cdot\Psi^{[a}\nabla^{b]}\nabla_c\phi\nabla^c\phi \,. \tag{A.2.5}$$

Here, we have introduced the coefficients Ω_i and Υ_i, which encodes the field dependence of the disformal coefficients $\tilde{C}$ and $\tilde{D}$ respectively. The subindex $_i$ indicates if the coefficient is generated through the derivative of ϕ or X. When there is no subindex, it means that Ω or Υ are sourced directly by the disformal coefficients. The disformal connection is not affected by the ϕ dependence of $\tilde{D}$. This becomes explicit by the absence of Υ_ϕ in (A.2.5). Also, the torsionless condition makes the transformed connection insensitive to the conformal factor alone. Thus, Ω does not appear either in (A.2.5).

Similarly, the disformal curvature 2-form can be computed through its definition

$$\begin{aligned}\tilde{\mathcal{R}}^{ab} =& \tilde{\mathcal{D}}\tilde{\omega}^{ab} = d\omega^{ab} + \omega^a{}_c \wedge \omega^{cb} + dX^{ab} + \omega^a{}_c \wedge X^{cb} + X^a{}_c \wedge \omega^{cb} + X^a{}_c \wedge X^{cb} \\ =& \mathcal{R}^{ab} + \mathcal{D}X^{ab} + X^a{}_c \wedge X^{cb}\end{aligned} \tag{A.2.6}$$

once $X^{ab} = \tilde{\omega}^{ab} - \omega^{ab}$ is known. Remarkably, this simple expression is valid for any field redefinition of the vielbein.

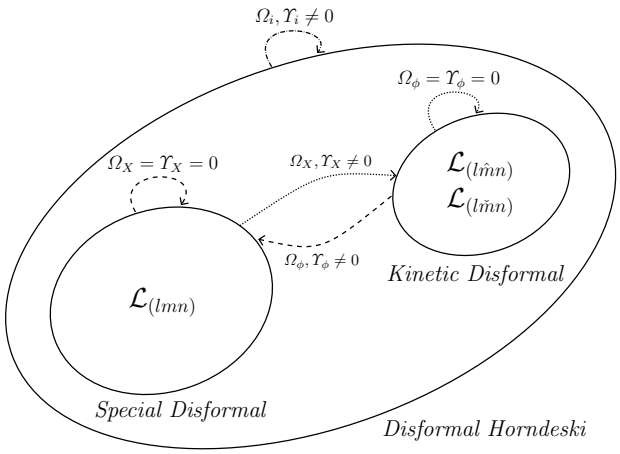

Figure A.2. Diagram of Horndeski's orbits. Each closed solid line represents a different orbit, which is defined as a set of theories invariant under a certain type of disformal transformations. The arrows indicate the connections of different orbits through specific field redefinitions defined by the coefficients Ω_i and Υ_i. There are three main orbits: the special disformal ($\Omega_X = \Upsilon_X = 0$), the kinetic disformal ($\Omega_\phi = \Upsilon_\phi = 0$) and the disformal Horndeski ($\Omega_i, \Upsilon_i \neq 0$). See more details in [3].

Onece we have transformed the building blocks, we can compute the disformal transformation of Horndeski theory. When the disformal transformation is special, i.e. when there is no dependence on X in the disformal coefficients and thus $\Omega_X = \Upsilon_X = 0$, we obtain that the transformed theory is also Horndeski, as it was first found in Ref. [95]. This result can be extended to any Lagrangian of the form $\mathcal{L}_{(lmn)}$ since, for special disformal transformations, the building blocks do not introduce new elements. Therefore, the set $\mathcal{L}_{(lmn)}$ could be defined as the set of Lagrangians invariant under special disformal transformations. Following this logic, we denote this set of Lagrangians as the *special disformal orbit*.

When the transformation is fully general, implying that there is a kinetic dependence in the disformal coefficients, new Lagrangians arise that are not part of the starting set. This defines the *first Horndeski's orbit*, i.e. the set of theories that are disformally related to Horndeski theory but which do not belong to it. In this section, we are going to compute the disformal transformation of this first Horndeski's orbit. Our objective is to find if a *second Horndeski's orbit* exits or the first one closes under disformal transformations. Of course, if one performs a disformal transformation on a disformal Horndeski theory, one will conclude that the final theory is also a disformal Horndeski theory. This is because disformal transformations can be composed so that one could define two consecutive transformations as only one with new coefficients. However, it is not a trivial question what the disformal transformation of the extended basis is. We find that the disformal transformation of the extended basis of Lagrangians will not generate new terms outside of the extended and the original set. Consequently, the first Horndeski's orbit closes under disformal transformations, meaning that there is only one

disformal Horndeski's orbit. This allows us to classify the different sets of Lagrangians, constructed with our building blocks, with respect to their invariance under disformal transformations in the following manner:

(a) *Special disformal orbit:* set of Lagrangians that is invariant under special disformal transformations, i.e. $\Omega_X = \Upsilon_X = 0$. In four dimensions, this set corresponds to Horndeski theory, which is built with the Lagrangians $\mathcal{L}_{(lmn)}$.

(b) *Kinetic disformal orbit:* set of Lagrangians that is invariant under kinetic disformal transformations, i.e. $\Omega_\phi = \Upsilon_\phi = 0$. In four dimensions, it is composed of $\mathcal{L}_{(0\hat{1}0)}$, $\mathcal{L}_{(0\hat{1}1)}$, $\mathcal{L}_{(0\hat{2}0)}$, $\mathcal{L}_{(0\hat{2}1)}$, $\mathcal{L}_{(0\check{1}0)}$ and $\mathcal{L}_{(0\hat{1}1)}$.

(c) *Disformal Horndeski's orbit:* set of Lagrangians that is invariant under disformal transformations. It is formed by the special and kinetic disformal orbits. It defines also the group of Lagrangians that can be disformally related to Horndeski theory.

This classification is also schematically presented in Fig. A.2.

Extended Disformal Transformations

We also consider generalized disformal transformation, containing second derivative, of the form

$$\begin{aligned}\tilde{\theta}^a =& \tilde{C}\theta^a + \tilde{D}_{nm}\left(\Psi^{mn}\right)^a \\ =& \tilde{C}\theta^a + \tilde{D}_{00}\Psi^a + \tilde{D}_{01}\bar{\Phi}^a + \tilde{D}_{10}\bar{\Psi}^a + \tilde{D}_{11}\hat{\Psi}^a + \mathcal{O}(n, m > 1)\,,\end{aligned} \tag{A.2.7}$$

where in the first line we are assuming Einstein's summation convention. These field redefinitions can be generically denoted as *extended disformal transformations*. One should notice that the coefficient $\tilde{D}_{00}$ represents the previous disformal coefficient $\tilde{D}$. Also, every $\tilde{D}_{nm}$ coefficient will have a different mass dimension to compensate the extra powers of second derivatives. Subsequently, we can repeat the process of finding an inverse. The main difference is that when contracting two Ψ^{nm} one does *not* get a term of higher order in second derivatives. Instead, one obtains $\left(\Psi^{mn}\right)^a{}_b\left(\Psi^{pq}\right)^b = \langle\Phi^{n+p}\rangle\left(\Psi^{mq}\right)^a$, which introduces an extra factor $\langle\Phi^{n+p}\rangle$. Thus, one can find an inverse. Nevertheless, its coefficients will be, in general, functions of scalars with n powers of second derivatives $\langle\Phi^n\rangle$. This means that this kind of transformations (A.2.7) represent *viable extensions* of the disformal transformations in which both the vielbein and its inverse are described with finite series of extended building blocks.

In addition to (A.2.7), we would like to discuss other possible routes to generalize disformal transformations. Specifically, one could consider adding more fields and with different spin:

Mixing with Spin-0 Fields. The simplest manner to enlarge a disformal transformation with extra fields would be to consider a transformation with N different scalar fields ϕ_A, where capital letters are indices of the internal field space $A = 1, \cdots, N$. Such a

transformation would read

$$\tilde{\theta}^a = \tilde{C}\theta^a + \sum_{A=1}^{N} \tilde{D}_A \left(\Psi_A\right)^a \,, \tag{A.2.8}$$

where $(\Psi_A)^a$ is the 1-form encoding the first derivatives associated to each field ϕ_A. The coefficients of this transformation would be functions of scalars built with the fields and their first derivatives. Applying this kind of transformations, one would end up with a multi-scalar-tensor theory. The main difference with respect to the single scalar case is that now there are many more possible interactions. In particular, an interaction between two first derivative 1-forms $(\Psi_A)^a \wedge (\Psi_B)^b$ is not zero if they are built with different fields, i.e. if $A \neq B$. This kind of new interaction arises already when transforming the volume element (or cosmological constant).

Mixing with Spin-1 Fields. More interestingly, one could mix fields with different spins as proposed in [96]. Restricting to integer spins, one could consider adding a spin-1 field to the vielbein. In analogy with the disformal transformation, one could define

$$\tilde{\theta}^a = \tilde{C}\theta^a + \tilde{V}\mathcal{A}^a \,, \tag{A.2.9}$$

where $\mathcal{A}^a = A^a\mathcal{A} = A^a A_b \theta^b$ is a 1-form encoding the vector field. Here, $\tilde{C}$ and $\tilde{V}$ would be functions of the modulus square of the vector field, which we parametrize for convenience with $-2X_V = A^a A_a$. Noticeably, one can recover a disformal transformation by going to the scalar limit $\mathcal{A}^a \to \Psi^a$. Consequently, we can benefit from all the machinery developed for disformal transformations. In this sense, we could introduce a 1-form encoding the first derivatives of the vector field $(\Phi_V)^a = \mathcal{D}A^a$, which would be analogous to the 1-form describing the second derivatives of the scalar field $\Phi^a = \mathcal{D}\nabla^a\phi$. With this dictionary in hand, one could translate all the previous calculations to obtain results in vector-tensor theories. In particular, if one starts from Lovelock's action and applies the vielbein redefinition (A.2.9) with constant coefficients, which is equivalent to a special disformal transformation, the resulting theory would correspond to the generalization of Proca theory in curved space [116]. The only precaution we must have in the transliteration is that now a covariant derivative does not conmute with the vector field, i.e. $(\Phi_V)^a = \mathcal{D}A^a \neq \nabla^a\mathcal{A}$, in contrast to the case of the gradient field, i.e. $\Phi^a = \mathcal{D}\nabla^a\phi = \nabla^a\mathcal{D}\phi$. Therefore, there would be additional terms with respect to Horndeski Lagrangian. These terms can be parametrized by the 1-form $\mathcal{F}^a = \nabla^a\mathcal{A} - \mathcal{D}A^a$, which can be related with the usual Abelian 2-form field strength $\mathcal{F} = \frac{1}{2}F_{ab}\theta^a\wedge\theta^b$ via an interior product.[4] With this additional building block, one could follow the same principles used in Ref. [1] for scalar-tensor theories to build a general vector-tensor theory.

In addition, one could think on combining both scalars and vectors fields. For the simplest case, this was studied in the context of TeVeS (see (1.1.13)), where the field

[4]Specifically, the 2-form field strength is given by $\mathcal{F} = \mathcal{D}\mathcal{A} = \frac{1}{2}(\nabla_a A_b - \nabla_b A_a)\theta^a \wedge \theta^b$. Thus, one realizes that $\mathcal{F}^a = \mathfrak{i}_{\delta^a}\mathcal{F}$, where $\delta^a \equiv \eta^{ab}\theta_b$ and $\mathfrak{i}_X$ is the interior product operator that maps p-forms onto $(p-1)$-forms by contracting the indices with the vector field X.

redefinition that generates the scalar-vector-tensor interactions reads

$$\tilde{\theta}^a = \tilde{C}(\phi)\theta^a + \tilde{V}(\phi)\mathcal{A}^a\,. \tag{A.2.10}$$

This represents a generalization of the simple vector-tensor transformation (A.2.9) in which the coefficients are allowed to depend on an additional scalar field ϕ. Another place where there is an interplay of spin-0, spin-1 and spin-2 fields is when a general dimensional reduction in the Kaluza-Klein framework is considered. In that case, the vector field arises from the non-diagonal components of the metric in the extra dimension.

Mixing with Spin-2 Fields. Finally, one could think on adding extra spin-2 fields. For simplicity, we will restrict to just one additional spin-2 field, but the following construction can be easily generalized to several spin-2 fields. In the language of differential forms, adding additional spin-2 fields means adding extra vielbeins [151]. We will denote the usual gravitational vielbein, to which matter is coupled, as before θ^a and the second dynamical vielbein will be encoded in Θ^a. Thus, a transformation that mixes both fields would be

$$\tilde{\theta}^a = \tilde{C}\theta^a + \tilde{F}\Theta^a\,, \tag{A.2.11}$$

where we set $\tilde{C}$ and $\tilde{F}$ to be constants. One can parametrize the transformed curvature as $\tilde{\mathcal{R}}^{ab} = \mathcal{R}^{ab} + \mathcal{R}_\Omega{}^{ab}$, where $\mathcal{R}_\Omega{}^{ab} = \mathcal{D}\Omega^{ab} + \Omega^a{}_c \wedge \Omega^{cb}$ contains the 2-form curvature associated to the second vielbein Θ^a and also derivative interactions of the two vielbeins, e.g. $\omega^a{}_c \wedge \Omega^{cb}$, where $\Omega^{ab} = \tilde{\omega}^{ab} - \omega^{ab}$ is the difference of the new connection and the original one.

CHAPTER

B

Phase lag test of c_g with LISA verification binaries

In theories in which matter is universally coupled to the metric, electromagnetic signals and ultrarelativistic particles propagate at the speed of light. If GWs have an anomalous propagation speed, $c_g \neq c$, this produces a delay between GW and electromagnetic signals

$$\Delta t = r\left(\frac{1}{c_g} - \frac{1}{c}\right) \equiv \frac{r}{c}\varepsilon_g \approx 10^{14} s \frac{r}{\text{Mpc}}\varepsilon_g \,, \tag{B.0.1}$$

where we define the *differential delay parameter* $\varepsilon_g \equiv c\partial\Delta t/\partial r$ (in general space-times r is the proper distance and one has to correct for time dilation at emission [17]). As we have discussed in section 3.4, the multi-messenger detection GW170817 has placed phenomenal constraints on the speed of GWs.

However, there are theories of gravity in which the speed of GWs varies with frequencies $c_g = c_g(k)$. Thus, it might happen that at $c_g(k_{\text{LIGO}}) = c$ but not at other frequencies. Then, one could aim at measuring the speed of GWs at lower frequencies with LISA using standard sirens. Nevertheless, remember that no distant GW-EM event will possibly be observed if $c_g(k_{\text{LISA}})$ is modified significantly, since the delay between both signals will be much larger than the monitoring time around the GW detection. Clearly, an alternative test for the speed of GWs would be needed to ensure a measurement of c_g at LISA frequencies. In the following, we discuss how observations of sources with periodic signals can help to test whether $c_g = c$. In particular, we propose a *phase lag test with eclipsing binaries* that overcomes this limitation.

The anomalous speed of GWs can be tested by monitoring periodic sources with both GW and EM emission [331,332]. This ensures that both signals can be observed continuously and allows for a long observation period. A suitable source is a binary system in the band of space-based interferometers [495], including *verification binaries* [496–498]: systems expected to be resolvable by LISA and which have already been identified and characterized using electromagnetic observations (see Ref. [498] for an updated list). An extraordinarily clean binary system is WDS J0651+2844: a binary, detached white dwarf system $\sim$ 1kpc away from the Sun and whose orbital plane is approximately aligned with the Solar System, allowing the observation of periodic eclipses [499]. Its short orbital period $\sim$ 12.75 min falls within the LISA band and makes it a loud GW source, in which

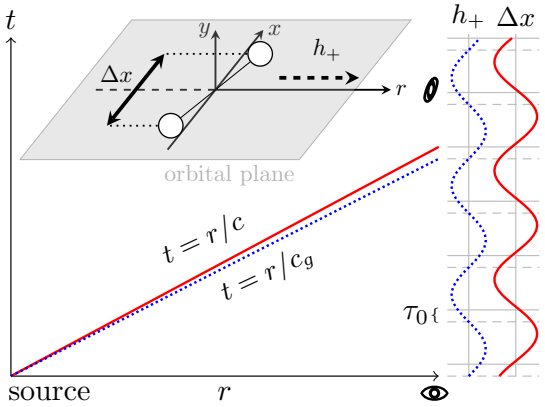

Figure B.1. The phase lag test for the speed of gravity. A compact binary system such as WDS J0651+2844 is monitored both electromagnetically and using GWs. For this geometry (top) only the + GW polarization is emitted in the observer's direction. Its amplitude h_+ is initially correlated with the object transverse separation Δx, but a phase lag (B.0.2) accumulates on the propagation if $c_g \neq c$ (bottom and right).

the effect of GW emission has already been observed by the period variation [500].

Let us model WDS J0651+2844 as a binary orbit coplanar with the observer and at a distance r from it, cf Fig. B.1. Due to symmetry the gravitational radiation emitted in the observer's direction will be predominantly in the + polarization $h_{ij} = h_+(t)(\hat{x}\hat{x} - \hat{y}\hat{y})$.[1] Assuming GR (i.e. $c_g = c$), the h_+ polarization will be in phase with Δx, the distance between the objects transverse to the line of sight as observed electromagnetically. Therefore, although the components of the binary will not be resolvable, $\Delta x = 0$ coincides with the eclipses and can be timed with extraordinary precision [331].

In theories other than GR, the EM and GW observables will evolve as periodic functions of different retarded times, i.e. $\Delta x \propto \cos(2\omega(t - r/c))$ and $h_+ \propto \cos(2\omega(t - r/c_g))$. The difference in propagation speed accumulated over the propagation distance r produces a *phase lag* between the GW and the EM signals:[2]

$$\Delta\Phi(t) \quad = \quad 2\omega\frac{r(t)}{c}\left(\frac{c}{c_g} - 1\right) = 2\omega\frac{r(t)}{c}\varepsilon_g\,, \tag{B.0.2}$$

where the distance between source and detector

$$r(t) = r_0 + v_{\rm rel}\,t + r_{\rm orb}(t)\,, \tag{B.0.3}$$

includes the initial separation, relative velocity and the detector's orbit. We will focus on the effect of $r_0, v_{\rm rel}$, as the effect of $r_{\rm orb}$ has been considered [333].

[1]The orbital inclination is $\iota = 86.9^{+1.6}_{-1.0}$ degrees [499], making $h_\times$ suppressed by $\cos(\iota) \approx 0.05$ in amplitude and shifted $\pi/2$ in phase relative to the + component.

[2]We have neglected the delay from the atmospheric or interstellar refractive index, which can be shown to be unimportant [331].

For eclipsing binaries we can neglect the error in EM measurements in constructing the relative phase (B.0.2) $\Delta\Phi(t) \equiv 2\omega(\tau_0 + \hat{\beta}t)$. The precision will be then limited by our knowledge of the GW signal. We can obtain an estimate of fo the 1-σ uncertainties using the Fisher matrix formalism [501] for the following quantities:

$$\tau_0 \equiv \varepsilon_g \frac{r_0}{c}\,,\ \Delta\tau_0 = \frac{1}{\sqrt{2}\,\omega\,\Sigma} \approx 0.2s \left(\frac{2\pi/\omega}{765s}\right)\left(\frac{T}{5y}\right), \tag{B.0.4}$$

$$\hat{\beta} \equiv \varepsilon_g \frac{v_{\rm rel}}{c}\,,\ \Delta\hat{\beta} = \frac{\sqrt{3/2}}{\omega T \Sigma} \approx 10^{-8} \left(\frac{2\pi/\omega}{765s}\right)\left(\frac{T}{5y}\right), \tag{B.0.5}$$

where T is the observation time and Σ denotes the total signal-to-noise ratio of the GW detection (see Appendix). The expected detection significance of verification binaries with LISA is $\Sigma \sim 100\left(\frac{T}{1y}\right)$ [333].

A non-zero measurement of either (B.0.4, B.0.5) represents a smoking gun for $c_g \neq c$:

- τ_0*:* The relative phase of the signals can detect an anomalous propagation speed in the range $|\varepsilon_g| \gtrsim 2 \cdot 10^{-12}\left(\frac{\rm kpc}{r_0}\right)\left(\frac{\Delta\tau_0}{0.2s}\right)$. The false-negative case where $2r_0\varepsilon_g\omega/(c\pi)$ equals an integer within the measurement error is very unlikely (prob. $\approx \Sigma^{-1} \sim 0.2\%$) and can be excluded by observing multiple systems or measuring the frequency shift $\hat{\beta}$.

- $\hat{\beta}$*:* The relative velocity of the system induces a frequency shift, sensitive to anomalous GW speeds in the range $|\varepsilon_g| \gtrsim 10^{-4}\left(\frac{30{\rm km}/s}{v_{\rm rel}}\right)\left(\frac{\Delta\beta}{10^{-8}}\right)$. Despite the $(\omega T)^{-1}$ gain when observing over many cycles, this test is less competitive due to the non-relativistic factor.

Note that both the measurement of the relative phase and the velocity can be used as a test of $\varepsilon_g \neq 0$ *and* as a measurement of c_g. The latter application requires a measurement of either r_0 or $v_{\rm rel}$, which will almost certainly dominate the error . Nevertheless, clean systems such as WDS J0651+2844 will be able to confirm deviations from $c_g = c$ at the level of few parts in a trillion.

Signal to noise estimates

The signal-to-noise ratio Σ for a GW detection is given by

$$\Sigma^2 = \frac{1}{\sigma_f^2}\int_0^T \tilde{R}^2(t)dt \equiv \varrho\,. \tag{B.0.6}$$

Here $\tilde{R}$ is the response of the detector to the signal and σ_f^2 is the noise power at the GW frequency. We assume the GW to be monochromatic and follow Ref. [333] (see Ref. [501] for further details and cautionary notes). For a given detector the response function depends on the GW polarizations as $\tilde{R}(t) = A_+(t)h_+ + A_\times(t)h_\times$ where A_i contain information of the antenna pattern of the detector and its orientation as a function of time. However, as discussed in the text, we will consider the situation in

which only one polarization is received and assume that the errors in the electromagnetic signal are negligible. Therefore we can reconstruct the relative phase (Eq. B.0.2 in main text) directly

$$\tilde{R}(t) = \Upsilon \cos(\varpi t + \psi)\,, \tag{B.0.7}$$

where the signal has an overall amplitude Υ, which will not directly affect the reconstruction of ψ and ω.

The Fisher matrix is then given as the derivative of Eq. (B.0.6) with respect to the model parameters

$$F_{ij} = \frac{2}{\sigma_f^2}\int_0^T \frac{\partial\tilde{R}}{\partial\theta_i}\frac{\partial\tilde{R}}{\partial\theta_j}dt\,, \tag{B.0.8}$$

where $\theta_i = (\Upsilon, \varpi, \psi)$ collectively denotes the unknown parameters of the signal. The error in the parameter θ_i assuming the other ones are perfectly known is $(F_{ii})^{-1/2}$, while the error in a parameter marginalized over the rest is $\sqrt{(F^{-1})_{ii}}$.

The Fisher matrix elements read

$$\begin{aligned}
F_{\Upsilon\Upsilon} &= \frac{2}{\sigma_f^2}\int \cos^2(\varpi t+\psi)dt = 2\varrho/\Upsilon^2\,,\\
F_{\Upsilon\varpi} &= \frac{2}{\sigma_f^2}\int -t\sin(\varpi t+\psi)\Upsilon\cos(\varpi t+\psi)dt \sim \text{osc.}\,,\\
F_{\Upsilon\psi} &= \frac{2}{\sigma_f^2}\int -\Upsilon\cos(\varpi t+\psi)\sin(\varpi t+\psi)dt \sim \text{osc.}\,,\\
F_{\varpi\varpi} &= \frac{2}{\sigma_f^2}\int \Upsilon^2 t^2\sin^2(\varpi t+\psi)dt = 2\varrho\frac{t^2}{3} + \text{osc.}\,,\\
F_{\varpi\psi} &= \frac{2}{\sigma_f^2}\int \Upsilon^2 t\sin^2(\varpi t+\psi)dt = \varrho t + \text{osc.}\,,\\
F_{\psi\psi} &= \frac{2}{\sigma_f^2}\int \Upsilon^2\sin^2(\varpi t+\psi)dt = 2\varrho + \text{osc.}\,,
\end{aligned}$$

where osc. denotes oscillatory terms that become negligible for $T \gg \varpi^{-1}$ and we have used $\varrho = \frac{\Upsilon^2}{2\sigma_f^2}T$. Since $F_{\Upsilon\varpi}$, $F_{\Upsilon\psi}$ do not build up with time, the amplitude is uncorrelated with the frequency and the phase. However, ϖ, ψ are correlated with one another. The Fisher matrix and its inverse for the (ϖ, ψ) subspace are

$$\hat{F} = \varrho\begin{pmatrix} \frac{2}{3}T^2 & T \\ T & 2\end{pmatrix}\,, \qquad \hat{F}^{-1} = \frac{1}{\varrho}\begin{pmatrix} \frac{6}{T^2} & -\frac{3}{T} \\ -\frac{3}{T} & 2\end{pmatrix}\,. \tag{B.0.9}$$

From which we read the errors in the phase and frequency

$$\Delta\psi = \frac{\sqrt{2}}{\Sigma}\,, \quad \Delta\varpi = \frac{\sqrt{6}}{T\cdot\Sigma}\,, \tag{B.0.10}$$

which translate straightforwardly into the results Eqs. (B.0.4, B.0.5).

CHAPTER

C

Solving stochastic inflation beyond slow-roll

In this appendix we complement the content of chapter 8 by providing further details in some of the computations. In particular we show how to compute the noise, C.1; give details on the full equations for the correlations functions, C.2; review stochastic SR inflation, C.3; and compute the statistics of PBH formation, C.4.

C.1 Computing the noise

In section 8.1 we presented the Hamilton-Jacobi formalism for stochastic inflation. A key ingredient was the appearece of noise terms induced by the short-wave modes. Here we detail the computation of the noise. For this calculation, we follow the work of Ref. [463]. The starting point is to apply the coarse-graining to the fields $\phi = \phi_s + \bar{\phi}$ and $\pi_\phi = (\pi_\phi)_s + \bar{\pi}_\phi$ through

$$\phi_s = -\int \frac{d^3k}{(2\pi)^{3/2}} W\left(\frac{k}{k_\sigma}\right)\left[a_{\vec{k}}\phi_{\vec{k}}\, e^{-i\vec{k}\vec{x}} + a^\dagger_{\vec{k}}\phi^\star_{\vec{k}}\, e^{i\vec{k}\vec{x}}\right], \tag{C.1.1}$$

$$(\pi_\phi)_s = -\int \frac{d^3k}{(2\pi)^{3/2}} W\left(\frac{k}{k_\sigma}\right)\left[a_{\vec{k}}\pi_{\vec{k}}\, e^{-i\vec{k}\vec{x}} + a^\dagger_{\vec{k}}\pi^\star_{\vec{k}}\, e^{i\vec{k}\vec{x}}\right], \tag{C.1.2}$$

where $W\left(k/k_\sigma\right)$ is a window function selecting the modes with $k \gg k_\sigma = \sigma aH$ for $\sigma \ll 1$. Then, the evolution equations lead to

$$\frac{d\bar{\phi}}{dN} = \bar{\pi}_\phi + (\pi_\phi)_s - \frac{d\phi_s}{dN}\,, \tag{C.1.3}$$

$$\frac{d\bar{\pi}_\phi}{dN} = -(3-\epsilon)\bar{\pi}_\phi - \frac{V_{,\phi}}{H^2} - \frac{d(\pi_\phi)_s}{dN} - (3-\epsilon)(\pi_\phi)_s - \left(\left(\frac{k}{aH}\right)^2 + \frac{m^2_{\rm eff}}{H^2}\right)\phi_s\,, \tag{C.1.4}$$

where we have linearized in the short wavelengths modes and $m_{\rm eff}$ is the time dependent effective mass that can be related to $V_{,\phi\phi}$ and ϵ_n, cf. (8.1.9). Subsequently, using the mode equations for ϕ_k and π_k (8.1.8-8.1.9), one can realize that when taking the time derivatives of ϕ_s and $(\pi_\phi)_s$ all the terms associated with them cancel in the equations of

motion but the one associated with the time derivatives of the window function. These remaining terms are thus the ones defining the noise

$$\xi_\phi(N) = -\int \frac{d^3k}{(2\pi)^{3/2}} \frac{d}{dN}\left[W\left(\frac{k}{k_\sigma}\right)\right]\left[a_{\vec{k}}\phi_{\vec{k}}\, e^{-i\vec{k}\vec{x}} + a^\dagger_{\vec{k}}\phi^\star_{\vec{k}}\, e^{i\vec{k}\vec{x}}\right], \tag{C.1.5}$$

$$\xi_\pi(N) = -\int \frac{d^3k}{(2\pi)^{3/2}} \frac{d}{dN}\left[W\left(\frac{k}{k_\sigma}\right)\right]\left[a_{\vec{k}}\pi_{\vec{k}}\, e^{-i\vec{k}\vec{x}} + a^\dagger_{\vec{k}}\pi^\star_{\vec{k}}\, e^{i\vec{k}\vec{x}}\right], \tag{C.1.6}$$

For a Heaviside window function, $W(k/k_\sigma) = \Theta(k/k_\sigma - 1)$, its derivative is a delta function $\frac{d}{dN}\left[W\left(\frac{k}{k_\sigma}\right)\right] = \delta(k - k_\sigma)$. This implies that the integrals can be solved. The noise is characterized by the two-point auto-correlations

$$\Xi_{AB}(N, N') \equiv \langle 0|\xi_A(N)\xi_B(N')|0\rangle = \frac{k^3_\sigma(N)}{2\pi^2}\frac{d\ln k_\sigma}{dN} A_{k_\sigma} B^\star_{k_\sigma} \cdot \delta(N - N'), \tag{C.1.7}$$

where A, B could be either ϕ or π. The fact that the noise is uncorrelated at different times (white/Markovian noise) is related to the choice of window function. Since the derivative of the Heaviside window function is a Dirac distribution in momentum space, the product of the two noises evaluated at different times will lead to $\delta(k - k_\sigma(N))\delta(k - k_\sigma(N')) = \delta(N - N')$. Having a Gaussian, white noise allows to connect the Langevin equations with the Fokker-Planck as discussed in the main text.

C.2 Correlation functions in detail

In Sec. 8.2, we have explained how to rewrite the problem of solving the Langevin equations in terms of a set of first order differential equations for the statistical moments $\langle \delta\phi^m \delta\pi^n \rangle$. In this appendix we provide details of some of the formulae we have used. Since we want to compute the n-point functions of the stochastic fluctuations, $\delta\phi = \bar{\phi} - \phi_c$ and $\delta\pi = \bar{\pi}_\phi - \pi_c$, we have to rewrite the drift vector as

$$D_\phi = -(\delta\pi + \pi_c), \tag{C.2.1}$$

$$D_\pi = -\left(3 - \frac{\kappa^2}{2}(\delta\pi + \pi_c)^2\right)(\delta\pi + \pi_c) + \left(3 - \frac{\kappa^2}{2}(\delta\pi + \pi_c)^2\right)\frac{(\ln V)_{,\phi}|_{\delta\phi+\phi_c}}{\kappa^2}. \tag{C.2.2}$$

If we substitute these expression into the evolution equations (8.2.7), we obtain

$$\begin{aligned}\frac{d\langle\delta\phi^n\delta\pi^m\rangle}{dN} &= -n\langle\delta\phi^{n-1}\delta\pi^{m+1}\rangle - 3m\langle\delta\phi^n\delta\pi^m\rangle \\ &+ m\frac{\kappa^2}{2}\left(\langle\delta\phi^n\delta\pi^{m+2}\rangle + 3\langle\delta\phi^n\delta\pi^{m+1}\rangle\pi_c + 3\langle\delta\phi^n\delta\pi^m\rangle\pi_c^2\right) \\ &+ m\frac{(3-\epsilon)}{\kappa^2}\sum_{i=1}\frac{(\ln V)^{(i+1)}}{i!}\langle\delta\phi^{n+i}\delta\pi^{m-1}\rangle \\ &- \frac{m}{2}\left(\langle\delta\phi^n\delta\pi^{m+1}(\ln V)_{,\phi}\rangle + 2\langle\delta\phi^n\delta\pi^m(\ln V)_{,\phi}\rangle\pi_c\right) \\ &+ \frac{1}{2}n(n-1)\Xi_{\phi\phi}\langle\delta\phi^{n-2}\delta\pi^m\rangle + \frac{1}{2}m(m-1)\Xi_{\pi\pi}\langle\delta\phi^n\delta\pi^{m-2}\rangle \\ &+ \frac{1}{2}n\,m\left(\Xi_{\phi\pi}+\Xi_{\pi\phi}\right)\langle\delta\phi^{n-1}\delta\pi^{m-1}\rangle\,.\end{aligned} \tag{C.2.3}$$

Note that in the third line, we have explicitly written the sum over the expansion of $(\ln V)_{,\phi}$ starting at 1 since the zeroth order was cancelled by the D_π term in (8.2.7). In analogy, we could expand the other contributions of the potential as

$$\langle\delta\phi^m\delta\pi^n(\ln V)_{,\phi}\rangle = \sum_{i=0}\frac{(\ln V)^{(i+1)}}{i!}\langle\delta\phi^{m+i}\delta\pi^n\rangle\,, \tag{C.2.4}$$

where the sum begins at zeroth order.

For illustrative purpose, we have shown in Fig. 8.3 the comparison between a Gaussian and a generalized PDF

$$P(x) = \frac{1}{2^{\frac{3-k_4}{2-k_4}}\Gamma\left[\frac{3-k_4}{2-k_4}\right]\sigma}\left(1+\mathrm{Erf}\left[s_3\frac{x}{\sqrt{2\sigma^2}}\right]\right)\exp\left[-\frac{1}{2}\left|\frac{x}{\sigma}\right|^{2-k_4}\right]\,, \tag{C.2.5}$$

where s_3 and k_4 parametrize the skewness and kurtosis respectively. Notice that for $s_3 = k_4 = 0$ one recovers a normal distribution.

C.3 Stochastic slow-roll inflation

Within slow-roll inflation, the stochastic dynamics of the inflaton are governed by a single Langevin equation

$$\frac{d\phi}{dN} = D_\phi + \xi_\phi\,. \tag{C.3.1}$$

Using the Fokker-Planck equation we can compute the evolution for any n-point correlation

$$\frac{d\langle\delta\phi^n\rangle}{dN} = n\langle\delta\phi^{n-1}D_\phi\rangle - n\langle\delta\phi^{n-1}\rangle D_\phi + \frac{1}{2}n(n-1)\langle\delta\phi^{n-2}\rangle\Xi_{\phi\phi} \tag{C.3.2}$$

$$= n\sum_{i=0}^{p-n}\frac{D_\phi^{(i+1)}}{(i+1)!}\langle\delta\phi^{n+i}\rangle + \frac{1}{2}n(n-1)\langle\delta\phi^{n-2}\rangle\Xi_{\phi\phi}\,, \tag{C.3.3}$$

where p is the maximum order considered. If D_ϕ is a polynomial, p is determined by its order. Otherwise, we have to truncate the serie at a given order.

If we want to compute the 2-point correlation to its lower order, we have to solve

$$\frac{d\langle\delta\phi^2\rangle}{dN} = 2D'_\phi\langle\delta\phi^2\rangle + \Xi_{\phi\phi}\,. \tag{C.3.4}$$

Recalling that the classical solution is given by $d\phi/dN = D_\phi$, we have

$$\frac{d\langle\delta\phi^2\rangle}{dN} - \frac{2}{D_\phi}\frac{dD_\phi}{dN}\langle\delta\phi^2\rangle = \Xi_{\phi\phi}\,, \tag{C.3.5}$$

whose solution is

$$\langle\delta\phi^2\rangle = D_\phi^2\int\frac{\Xi_{\phi\phi}}{D_\phi^2}dN\,. \tag{C.3.6}$$

Using that $D_\phi^2 = 2\epsilon/\kappa^2$ one can show that

$$\frac{d\langle\delta\phi^2\rangle}{dN} - \epsilon_2\langle\delta\phi^2\rangle = \Xi_{\phi\phi}\,, \tag{C.3.7}$$

and therefore

$$\frac{d\zeta^2}{dN} = \frac{d}{dN}\left(\frac{\langle\delta\phi^2\rangle}{\phi'^2}\right) = \frac{\kappa^2\Xi_{\phi\phi}}{2\epsilon}\,. \tag{C.3.8}$$

Recalling that $\Xi_{\phi\phi} = (1-\epsilon)\mathcal{P}_{\delta\phi}$, then

$$\mathcal{P}_\zeta = \frac{d\langle\zeta(N)\rangle^2}{d\ln k} = \frac{1}{1-\epsilon}\frac{d}{dN}\left(\frac{\langle\delta\phi^2\rangle}{\phi'^2}\right) = \frac{\kappa^2\mathcal{P}_{\delta\phi}}{2\epsilon}\,, \tag{C.3.9}$$

which is the standard result for the power spectrum.

We can also compute the noise analytically from the mode equation, which reads

$$\frac{d^2\phi_k}{dN^2} + (3-\epsilon)\frac{d\phi_k}{dN} + \left(\frac{V''(\phi)}{H^2} + \left(\frac{k}{aH}\right)^2\right)\phi_k = 0\,. \tag{C.3.10}$$

At leading order in SR, $H = \text{const} + \mathcal{O}(\epsilon)$, and for a free field, $V = m^2\phi^2/2$, there is an exact solution, in terms of the Hankel functions $H_\nu^{(1)}$,

$$\phi_k = \frac{\sqrt{\pi}}{2a}\left(\frac{1}{aH}\right)^{1/2} H_\nu^{(1)}\left[\frac{k}{aH}\right] \tag{C.3.11}$$

which can be approximated after horizon crossing to

$$|\phi_k| \to \frac{H}{\sqrt{2k^3}}\left(\frac{k}{aH}\right)^{\frac{3}{2}-\nu} \tag{C.3.12}$$

with $\nu^2 = 9/4 - m^2/H^2$. Thus,

$$\pi_k = \frac{d\phi_k}{dN} = \frac{H}{\sqrt{2k^3}}\left(\nu - \frac{3}{2}\right)\left(\frac{k}{aH}\right)^{\frac{3}{2}-\nu} + \mathcal{O}(\epsilon') \simeq -\frac{H}{\sqrt{2k^3}}\left(\frac{m^2}{3H^2}\right)\left(\frac{k}{aH}\right)^{\frac{3}{2}-\nu} \quad \text{(C.3.13)}$$

using that $da/dN = a$. Note also that $m^2/3H^2 = V''/\kappa^2 V \equiv \eta_{SR}$. This result shows that π_k is higher order in SR. Therefore, to leading order in SR only $\Xi_{\phi\phi}$ will contribute to the noise by (recall $k_\sigma = \sigma a H$)

$$\Xi_{\phi\phi}(N) = \frac{k_\sigma^3}{2\pi^2}\frac{d\ln k_\sigma}{dN}\frac{H^2}{2k_\sigma^3}\left(\frac{k_\sigma}{aH}\right)^{3-2\nu} = \frac{H^2}{4\pi^2}\sigma^{3-2\nu} \simeq \frac{H^2}{4\pi^2}\,, \quad \text{(C.3.14)}$$

where in the last equality we have used that at this order in SR $\sigma^{3-2\nu} = \sigma^{\mathcal{O}(\epsilon^2)} \simeq 1$. In total, we recover the standard Langevin equation

$$\frac{d\phi}{dN} = -\frac{V_{,\phi}}{\kappa^2 V} + \frac{H}{2\pi}\,\xi\,, \quad \text{(C.3.15)}$$

where the noise is normalized as $\langle \xi(N)\xi(N')\rangle = \delta(N - N')$.

C.4 Statistics of PBH formation

Due to its relevance, we summarize the derivation of the exact formula for $P(\zeta > \zeta_c)$ presented in Eq. (8.3.3). We start the derivation from the last equality of (8.3.2)

$$P(\zeta > \zeta_c) = \int_{\zeta_c}^{\infty} da \int_{-\infty}^{\infty} \frac{du}{2\pi} e^{-iua}\chi(u) = \int_{\zeta_c}^{\infty} da \int_{-\infty}^{\infty} \frac{du}{2\pi} \exp\left[\sum_{p=1}^{\infty} \frac{i^p}{p!}\kappa_p u_\phi^p\right] e^{-iua}\,, \quad \text{(C.4.1)}$$

where we have rewritten the characteristic function in terms of the cumulants following (8.2.22).

The first trick of the derivation is to rewrite all the dependence in u as Gaussian that we can later integrate. For that, one can use the simple identity

$$u e^{-iua} = +i\frac{\partial}{\partial a} e^{-iua} \quad \text{(C.4.2)}$$

to arrive at

$$P(\zeta > \zeta_c) = \int_{\zeta_c}^{\infty} da \int_{-\infty}^{\infty} \frac{du}{2\pi} \exp\left[\sum_{p=3}^{\infty} \frac{(-1)^p}{p!}\kappa_p \frac{\partial^p}{\partial a^p}\right] e^{-\frac{1}{2}\kappa_2 u^2 - i(a-\kappa_1)u}\,, \quad \text{(C.4.3)}$$

which can be integrated (using the standard change of variables) in u

$$P(\zeta > \zeta_c) = \frac{1}{\sqrt{2\pi\kappa_2}} \int_{\zeta_c}^{\infty} da \exp\left[\sum_{p=3}^{\infty} \frac{(-1)^p}{p!}\kappa_p \frac{\partial^p}{\partial a^p}\right] e^{-\frac{(a-\kappa_1)^2}{2\kappa_2}}\,. \quad \text{(C.4.4)}$$

The second trick of the derivation is to expand the exponential with the partial derivatives in terms of a unique series that separates the Gaussian and non-Gaussian contributions

$$\begin{aligned} P(\zeta > \zeta_c) &= \frac{1}{\sqrt{2\pi}} \int_{z_c}^{\infty} dz \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{p=3}^{\infty} \frac{(-1)^p}{p!} \bar{\kappa}_p \frac{\partial^p}{\partial z^p} \right)^n e^{-\frac{z^2}{2}} \\ &= \frac{1}{\sqrt{2\pi}} \int_{z_c}^{\infty} dz \left(1 + \sum_{n=3}^{\infty} \frac{(-1)^n}{n!} \bar{K}_n \frac{\partial^n}{\partial z^n} \right) e^{-\frac{z^2}{2}} . \end{aligned} \tag{C.4.5}$$

Here we have first changed variables to $z = (a - \kappa_1)/\kappa_2^{1/2}$, $z_c = (\zeta_c - \kappa_1)/\kappa_2^{1/2}$ and normalized $\bar{\kappa}_n = \kappa_n/\kappa_2^{n/2}$. Then, we have expanded the serie and defined new functions $\bar{K}_n$ that accounts for the different combinations of normalized cumulants. Note that $\bar{K}_n$ only starts to differ from $\bar{\kappa}_n$ at sixth order, when the terms from $n = 2$ of the first line of (C.4.5) appear. In this way, and recalling the definition of the complementary error function

$$\mathrm{Erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-z^2} dz \tag{C.4.6}$$

and of the Hermite polynomials

$$H_n(x) \equiv (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2} , \tag{C.4.7}$$

one can arrive to the final result (8.3.3). While the identification of the Erfc is direct from the first term of (C.4.5), the Hermite polynomials can be introduced by noting that

$$\int_{z_c}^{\infty} \frac{\partial^n}{\partial z^n} e^{-\frac{z^2}{2}} = \left[\frac{\partial^{n-1}}{\partial z^{n-1}} e^{-\frac{z^2}{2}} \right]_{z=z_c}^{z=\infty} = \frac{e^{-\frac{z_c^2}{2}}}{(-1)^n 2^{(n-1)/2}} H_{n-1}\left[\frac{z_c}{\sqrt{2}} \right] , \tag{C.4.8}$$

since the contribution of the integral at infinity vanishes. With this, the derivation of (8.3.3) is completed.

Bibliography

[1] J. M. Ezquiaga, J. García-Bellido, and M. Zumalacárregui, *Towards the most general scalar-tensor theories of gravity: a unified approach in the language of differential forms*, *Phys. Rev.* **D94** (2016), no. 2 024005, [arXiv:1603.01269].

[2] D. Bettoni, J. M. Ezquiaga, K. Hinterbichler, and M. Zumalacárregui, *Speed of Gravitational Waves and the Fate of Scalar-Tensor Gravity*, *Phys. Rev.* **D95** (2017), no. 8 084029, [arXiv:1608.01982].

[3] J. M. Ezquiaga, J. García-Bellido, and M. Zumalacárregui, *Field redefinitions in theories beyond Einstein gravity using the language of differential forms*, *Phys. Rev.* **D95** (2017), no. 8 084039, [arXiv:1701.05476].

[4] J. M. Ezquiaga, J. Garcia-Bellido, and E. Ruiz Morales, *Primordial Black Hole production in Critical Higgs Inflation*, *Phys. Lett.* **B776** (2018) 345–349, [arXiv:1705.04861].

[5] J. M. Ezquiaga and M. Zumalacárregui, *Dark Energy After GW170817: Dead Ends and the Road Ahead*, *Phys. Rev. Lett.* **119** (2017), no. 25 251304, [arXiv:1710.05901].

[6] J. M. Ezquiaga and J. Garca-Bellido, *Quantum diffusion beyond slow-roll: implications for primordial black-hole production*, *JCAP* **1808** (2018) 018, [arXiv:1805.06731].

[7] J. M. Ezquiaga and M. Zumalacrregui, *Dark Energy in light of Multi-Messenger Gravitational-Wave astronomy*, *Front. Astron. Space Sci.* **5** (2018) 44, [arXiv:1807.09241].

[8] **LISA Cosmology Working Group** Collaboration, E. Belgacem et al., *Testing modified gravity at cosmological distances with LISA standard sirens*, arXiv:1906.01593.

[9] J. M. Ezquiaga and M. Zumalacrregui *In preparation.* (2019).

[10] J. Beltran Jimenez, J. M. Ezquiaga, and L. Heisenberg *In preparation.* (2019).

[11] J. M. Ezquiaga, J. Garcia-Bellido, and V. Vennin *In preparation.* (2019).

[12] **Planck** Collaboration, N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, arXiv:1807.06209.

[13] S. Weinberg, *The Cosmological Constant Problem*, *Rev. Mod. Phys.* **61** (1989) 1–23.

[14] J. Martin, *Everything You Always Wanted To Know About The Cosmological Constant Problem (But Were Afraid To Ask)*, *Comptes Rendus Physique* **13** (2012) 566–665, [arXiv:1205.3365].

[15] E. J. Copeland, M. Sami, and S. Tsujikawa, *Dynamics of dark energy*, *Int. J. Mod. Phys.* **D15** (2006) 1753–1936, [hep-th/0603057].

[16] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, *Modified Gravity and Cosmology*, *Phys. Rept.* **513** (2012) 1–189, [arXiv:1106.2476].

[17] C. M. Will, *The Confrontation between General Relativity and Experiment*, *Living Rev. Rel.* **17** (2014) 4, [arXiv:1403.7377].

[18] E. Berti et al., *Testing General Relativity with Present and Future Astrophysical Observations*, *Class. Quant. Grav.* **32** (2015) 243001, [arXiv:1501.07274].

[19] A. M. Archibald, N. V. Gusinskaia, J. W. T. Hessels, A. T. Deller, D. L. Kaplan, D. R. Lorimer, R. S. Lynch, S. M. Ransom, and I. H. Stairs, *Testing the universality of free fall by tracking a pulsar in a stellar triple system*, arXiv:1807.02059.

[20] A. Hees et al., *Testing General Relativity with stellar orbits around the supermassive black hole in our Galactic center*, *Phys. Rev. Lett.* **118** (2017), no. 21 211101, [arXiv:1705.07902].

[21] P. Brax, *Screening mechanisms in modified gravity*, *Class. Quant. Grav.* **30** (2013) 214005.

[22] A. Joyce, B. Jain, J. Khoury, and M. Trodden, *Beyond the Cosmological Standard Model*, *Phys. Rept.* **568** (2015) 1–98, [arXiv:1407.0059].

[23] D. H. Weinberg, M. J. Mortonson, D. J. Eisenstein, C. Hirata, A. G. Riess, and E. Rozo, *Observational Probes of Cosmic Acceleration*, *Phys. Rept.* **530** (2013) 87–255, [arXiv:1201.2434].

[24] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott et al., *GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs*, arXiv:1811.12907.

[25] **Virgo, LIGO Scientific** Collaboration, B. P. Abbott et al., *GW170817: Implications for the Stochastic Gravitational-Wave Background from Compact Binary Coalescences*, *Phys. Rev. Lett.* **120** (2018), no. 9 091101, [arXiv:1710.05837].

[26] **Virgo, LIGO Scientific** Collaboration, B. P. Abbott et al., *Full Band All-sky Search for Periodic Gravitational Waves in the O1 LIGO Data*, *Phys. Rev.* **D97** (2018), no. 10 102003, [arXiv:1802.05241].

[27] **Virgo, LIGO Scientific** Collaboration, B. P. Abbott et al., *All-sky search for long-duration gravitational wave transients in the first Advanced LIGO observing run*, *Class. Quant. Grav.* **35** (2018), no. 6 065009, [arXiv:1711.06843].

[28] M. Sasaki, T. Suyama, T. Tanaka, and S. Yokoyama, *Primordial black holesperspectives in gravitational wave astronomy*, *Class. Quant. Grav.* **35** (2018), no. 6 063001, [arXiv:1801.05235].

[29] D. Blas, D. L. Nacir, and S. Sibiryakov, *Ultralight Dark Matter Resonates with Binary Pulsars*, *Phys. Rev. Lett.* **118** (2017), no. 26 261102, [arXiv:1612.06789].

[30] A. Arvanitaki, M. Baryakhtar, S. Dimopoulos, S. Dubovsky, and R. Lasenby, *Black Hole Mergers and the QCD Axion at Advanced LIGO*, *Phys. Rev.* **D95** (2017), no. 4 043001, [arXiv:1604.03958].

[31] **Virgo, Fermi-GBM, INTEGRAL, LIGO Scientific** Collaboration, B. P. Abbott et al., *Gravitational Waves and Gamma-rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A*, *Astrophys. J.* **848** (2017), no. 2 L13, [arXiv:1710.05834].

[32] B. P. Abbott et al., *Multi-messenger Observations of a Binary Neutron Star Merger*, *Astrophys. J.* **848** (2017), no. 2 L12, [arXiv:1710.05833].

[33] **LIGO Scientific, VINROUGE, Las Cumbres Observatory, DES, DLT40, Virgo, 1M2H, Dark Energy Camera GW-E, MASTER** Collaboration, B. P. Abbott et al., *A gravitational-wave standard siren measurement of the Hubble constant*, *Nature* **551** (2017), no. 7678 85–88, [arXiv:1710.05835].

[34] **DES, LIGO Scientific, Virgo** Collaboration, M. Soares-Santos et al., *First Measurement of the Hubble Constant from a Dark Standard Siren using the Dark Energy Survey Galaxies and the LIGO/Virgo BinaryBlack-hole Merger GW170814*, *Astrophys. J.* **876** (2019), no. 1 L7, [arXiv:1901.01540].

[35] S. Weinberg, *Photons and Gravitons in s Matrix Theory: Derivation of Charge Conservation and Equality of Gravitational and Inertial Mass*, *Phys. Rev.* **135** (1964) B1049–B1056.

[36] S. Weinberg, *Photons and gravitons in perturbation theory: Derivation of Maxwell's and Einstein's equations*, *Phys. Rev.* **138** (1965) B988–B1002.

[37] J. J. van der Bij, H. van Dam, and Y. J. Ng, *The Exchange of Massless Spin Two Particles*, *Physica* **116A** (1982) 307–320.

[38] D. Lovelock, *The Einstein tensor and its generalizations*, *J. Math. Phys.* **12** (1971) 498–501.

[39] D. Lovelock, *The four-dimensionality of space and the einstein tensor*, *J. Math. Phys.* **13** (1972) 874–876.

[40] D. Bailin and A. Love, *KALUZA-KLEIN THEORIES*, *Rept. Prog. Phys.* **50** (1987) 1087–1170.

[41] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, *The Hierarchy problem and new dimensions at a millimeter*, *Phys. Lett.* **B429** (1998) 263–272, [hep-ph/9803315].

[42] G. R. Dvali, G. Gabadadze, and M. Porrati, *4-D gravity on a brane in 5-D Minkowski space*, *Phys. Lett.* **B485** (2000) 208–214, [hep-th/0005016].

[43] A. Nicolis and R. Rattazzi, *Classical and quantum consistency of the DGP model*, *JHEP* **06** (2004) 059, [hep-th/0404159].

[44] A. Nicolis, R. Rattazzi, and E. Trincherini, *The Galileon as a local modification of gravity*, *Phys. Rev.* **D79** (2009) 064036, [arXiv:0811.2197].

[45] D. Blas and E. Lim, *Phenomenology of theories of gravity without Lorentz invariance: the preferred frame case*, *Int. J. Mod. Phys.* **D23** (2015) 1443009, [arXiv:1412.4828].

[46] P. Horava, *Quantum Gravity at a Lifshitz Point*, *Phys. Rev.* **D79** (2009) 084008, [arXiv:0901.3775].

[47] T. Jacobson, *Einstein-aether gravity: A Status report*, *PoS* **QG-PH** (2007) 020, [arXiv:0801.1547].

[48] D. Blas, O. Pujolas, and S. Sibiryakov, *Models of non-relativistic quantum gravity: The Good, the bad and the healthy*, *JHEP* **04** (2011) 018, [arXiv:1007.3503].

[49] T. Jacobson, *Extended Horava gravity and Einstein-aether theory*, *Phys. Rev.* **D81** (2010) 101502, [arXiv:1001.4823]. [Erratum: Phys. Rev.D82,129901(2010)].

[50] B. Audren, D. Blas, J. Lesgourgues, and S. Sibiryakov, *Cosmological constraints on Lorentz violating dark energy*, *JCAP* **1308** (2013) 039, [arXiv:1305.0009].

[51] B. Audren, D. Blas, M. M. Ivanov, J. Lesgourgues, and S. Sibiryakov, *Cosmological constraints on deviations from Lorentz invariance in gravity and dark matter*, *JCAP* **1503** (2015), no. 03 016, [arXiv:1410.6514].

[52] S. Deser and R. P. Woodard, *Nonlocal Cosmology*, *Phys. Rev. Lett.* **99** (2007) 111301, [arXiv:0706.2151].

[53] T. S. Koivisto, *Newtonian limit of nonlocal cosmology*, *Phys. Rev.* **D78** (2008) 123505, [arXiv:0807.3778].

[54] M. Jaccard, M. Maggiore, and E. Mitsou, *Nonlocal theory of massive gravity*, *Phys. Rev.* **D88** (2013), no. 4 044033, [arXiv:1305.3034].

[55] M. Maggiore, *Phantom dark energy from nonlocal infrared modifications of general relativity*, *Phys. Rev.* **D89** (2014), no. 4 043008, [arXiv:1307.3898].

[56] M. Maggiore and M. Mancarella, *Nonlocal gravity and dark energy*, *Phys. Rev.* **D90** (2014), no. 2 023005, [arXiv:1402.0448].

[57] A. Barreira, B. Li, W. A. Hellwing, C. M. Baugh, and S. Pascoli, *Nonlinear structure formation in Nonlocal Gravity*, *JCAP* **1409** (2014), no. 09 031, [arXiv:1408.1084].

[58] G. Calcagni and L. Modesto, *Nonlocal quantum gravity and M-theory*, *Phys. Rev.* **D91** (2015), no. 12 124059, [arXiv:1404.2137].

[59] P. G. Ferreira and A. L. Maroto, *A few cosmological implications of tensor nonlocalities*, *Phys. Rev.* **D88** (2013), no. 12 123502, [arXiv:1310.1238].

[60] H. Nersisyan, Y. Akrami, L. Amendola, T. S. Koivisto, J. Rubio, and A. R. Solomon, *Instabilities in tensorial nonlocal gravity*, *Phys. Rev.* **D95** (2017), no. 4 043539, [arXiv:1610.01799].

[61] L. Heisenberg, *A systematic approach to generalisations of General Relativity and their cosmological implications*, *Phys. Rept.* **796** (2019) 1–113, [arXiv:1807.01725].

[62] C. de Rham and A. J. Tolley, *DBI and the Galileon reunited*, *JCAP* **1005** (2010) 015, [arXiv:1003.5917].

[63] G. Goon, K. Hinterbichler, and M. Trodden, *A New Class of Effective Field Theories from Embedded Branes*, *Phys. Rev. Lett.* **106** (2011) 231102, [arXiv:1103.6029].

[64] T. Koivisto, D. Wills, and I. Zavala, *Dark D-brane Cosmology*, *JCAP* **1406** (2014) 036, [arXiv:1312.2597].

[65] G. W. Horndeski, *Second-order scalar-tensor field equations in a four-dimensional space*, *Int. J. Theor. Phys.* **10** (1974) 363–384.

[66] M. Ostrogradski *Mem. Ac. St. Petersbourg VI* **4** (1850) 385.

[67] R. P. Woodard, *Ostrogradsky's theorem on Hamiltonian instability*, *Scholarpedia* **10** (2015), no. 8 32243, [arXiv:1506.02210].

[68] J. Z. Simon, *Higher Derivative Lagrangians, Nonlocality, Problems and Solutions*, *Phys.Rev.* **D41** (1990) 3720.

[69] H. Motohashi, K. Noui, T. Suyama, M. Yamaguchi, and D. Langlois, *Healthy degenerate theories with higher derivatives*, *JCAP* **1607** (2016), no. 07 033, [arXiv:1603.09355].

[70] D. Langlois and K. Noui, *Degenerate higher derivative theories beyond Horndeski: evading the Ostrogradski instability*, *JCAP* **1602** (2016), no. 02 034, [arXiv:1510.06930].

[71] H. Motohashi, T. Suyama, and M. Yamaguchi, *Ghost-free theories with arbitrary higher-order time derivatives*, *JHEP* **06** (2018) 133, [arXiv:1804.07990].

[72] C. Brans and R. H. Dicke, *Mach's principle and a relativistic theory of gravitation*, *Phys. Rev.* **124** (1961) 925–935.

[73] C. Wetterich, *Cosmology and the Fate of Dilatation Symmetry*, *Nucl.Phys.* **B302** (1988) 668.

[74] B. Ratra and P. J. E. Peebles, *Cosmological Consequences of a Rolling Homogeneous Scalar Field*, *Phys. Rev.* **D37** (1988) 3406.

[75] C. Armendariz-Picon, T. Damour, and V. F. Mukhanov, *k-Inflation*, *Phys. Lett.* **B458** (1999) 209–218, [hep-th/9904075].

[76] C. Armendariz-Picon, V. F. Mukhanov, and P. J. Steinhardt, *Essentials of k essence*, *Phys. Rev.* **D63** (2001) 103510, [astro-ph/0006373].

[77] S. M. Carroll, V. Duvvuri, M. Trodden, and M. S. Turner, *Is cosmic speed - up due to new gravitational physics?*, *Phys.Rev.* **D70** (2004) 043528, [astro-ph/0306438].

[78] W. Hu and I. Sawicki, *Models of f(R) Cosmic Acceleration that Evade Solar-System Tests*, *Phys. Rev.* **D76** (2007) 064004, [arXiv:0705.1158].

[79] T. P. Sotiriou and V. Faraoni, *f(R) Theories Of Gravity*, *Rev. Mod. Phys.* **82** (2010) 451–497, [arXiv:0805.1726].

[80] A. De Felice and S. Tsujikawa, *f(R) theories*, *Living Rev. Rel.* **13** (2010) 3, [arXiv:1002.4928].

[81] J. Khoury and A. Weltman, *Chameleon fields: Awaiting surprises for tests of gravity in space*, *Phys.Rev.Lett.* **93** (2004) 171104.

[82] K. Hinterbichler and J. Khoury, *Symmetron Fields: Screening Long-Range Forces Through Local Symmetry Restoration*, *Phys.Rev.Lett.* **104** (2010) 231301, [arXiv:1001.4525].

[83] E. E. Flanagan, *The Conformal frame freedom in theories of gravitation*, *Class. Quant. Grav.* **21** (2004) 3817, [gr-qc/0403063].

[84] A. Avilez and C. Skordis, *Cosmological constraints on Brans-Dicke theory*, *Phys. Rev. Lett.* **113** (2014), no. 1 011101, [arXiv:1303.4330].

[85] C. Deffayet, X. Gao, D. A. Steer, and G. Zahariade, *From k-essence to generalised Galileons*, *Phys. Rev.* **D84** (2011) 064039, [arXiv:1103.3260].

[86] T. Kobayashi, M. Yamaguchi, and J. Yokoyama, *Generalized G-inflation: Inflation with the most general second-order field equations*, *Prog. Theor. Phys.* **126** (2011) 511–529, [arXiv:1105.5723].

[87] C. Deffayet, O. Pujolas, I. Sawicki, and A. Vikman, *Imperfect Dark Energy from Kinetic Gravity Braiding*, *JCAP* **1010** (2010) 026, [arXiv:1008.0048].

[88] T. Kobayashi, M. Yamaguchi, and J. Yokoyama, *G-inflation: Inflation driven by the Galileon field*, *Phys.Rev.Lett.* **105** (2010) 231302, [arXiv:1008.0603].

[89] O. Pujolas, I. Sawicki, and A. Vikman, *The Imperfect Fluid behind Kinetic Gravity Braiding*, *JHEP* **1111** (2011) 156, [arXiv:1103.5360].

[90] C. Deffayet, G. Esposito-Farese, and A. Vikman, *Covariant Galileon*, *Phys. Rev.* **D79** (2009) 084003, [arXiv:0901.1314].

[91] T. S. Koivisto, D. F. Mota, and M. Zumalacarregui, *Screening Modifications of Gravity through Disformally Coupled Fields*, *Phys. Rev. Lett.* **109** (2012) 241102, [arXiv:1205.3167].

[92] M. Zumalacarregui, T. S. Koivisto, and D. F. Mota, *DBI Galileons in the Einstein Frame: Local Gravity and Cosmology*, *Phys. Rev.* **D87** (2013) 083010, [arXiv:1210.8016].

[93] C. Charmousis, E. J. Copeland, A. Padilla, and P. M. Saffin, *General second order scalar-tensor theory, self tuning, and the Fab Four*, *Phys. Rev. Lett.* **108** (2012) 051101, [arXiv:1106.2000].

[94] P. Martin-Moruno, N. J. Nunes, and F. S. N. Lobo, *Horndeski theories self-tuning to a de Sitter vacuum*, *Phys. Rev.* **D91** (2015), no. 8 084029, [arXiv:1502.03236].

[95] D. Bettoni and S. Liberati, *Disformal invariance of second order scalar-tensor theories: Framing the Horndeski action*, *Phys. Rev.* **D88** (2013) 084020, [arXiv:1306.6724].

[96] M. Zumalacárregui and J. García-Bellido, *Transforming gravity: from derivative couplings to matter to second-order scalar-tensor theories beyond the Horndeski Lagrangian*, *Phys. Rev.* **D89** (2014) 064046, [arXiv:1308.4685].

[97] J. D. Bekenstein, *The Relation between physical and gravitational geometry*, *Phys. Rev.* **D48** (1993) 3641–3647, [gr-qc/9211017].

[98] A. H. Chamseddine and V. Mukhanov, *Mimetic Dark Matter*, *JHEP* **11** (2013) 135, [arXiv:1308.5410].

[99] P. Creminelli and F. Vernizzi, *Dark Energy after GW170817 and GRB170817A*, *Phys. Rev. Lett.* **119** (2017), no. 25 251302, [arXiv:1710.05877].

[100] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, *Healthy theories beyond Horndeski*, *Phys. Rev. Lett.* **114** (2015), no. 21 211101, [arXiv:1404.6495].

[101] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, *Exploring gravitational theories beyond Horndeski*, *JCAP* **1502** (2015) 018, [arXiv:1408.1952].

[102] C. Deffayet, G. Esposito-Farese, and D. A. Steer, *Counting the degrees of freedom of generalized Galileons*, *Phys. Rev.* **D92** (2015) 084013, [arXiv:1506.01974].

[103] M. Crisostomi, M. Hull, K. Koyama, and G. Tasinato, *Horndeski: beyond, or not beyond?*, *JCAP* **1603** (2016), no. 03 038, [arXiv:1601.04658].

[104] M. Crisostomi, K. Koyama, and G. Tasinato, *Extended Scalar-Tensor Theories of Gravity*, *JCAP* **1604** (2016), no. 04 044, [arXiv:1602.03119].

[105] J. Ben Achour, D. Langlois, and K. Noui, *Degenerate higher order scalar-tensor theories beyond Horndeski and disformal transformations*, *Phys. Rev.* **D93** (2016), no. 12 124005, [arXiv:1602.08398].

[106] M. Crisostomi, R. Klein, and D. Roest, *Higher Derivative Field Theories: Degeneracy Conditions and Classes*, *JHEP* **06** (2017) 124, [arXiv:1703.01623].

[107] J. Ben Achour, M. Crisostomi, K. Koyama, D. Langlois, K. Noui, and G. Tasinato, *Degenerate higher order scalar-tensor theories beyond Horndeski up to cubic order*, *JHEP* **12** (2016) 100, [arXiv:1608.08135].

[108] D. Langlois, M. Mancarella, K. Noui, and F. Vernizzi, *Effective Description of Higher-Order Scalar-Tensor Theories*, *JCAP* **1705** (2017), no. 05 033, [arXiv:1703.03797].

[109] J. A. R. Cembranos, C. Hallabrin, A. L. Maroto, and S. J. N. Jareno, *Isotropy theorem for cosmological vector fields*, *Phys. Rev.* **D86** (2012) 021301, [arXiv:1203.6221].

[110] J. A. R. Cembranos, A. L. Maroto, and S. J. Núñez Jareño, *Perturbations of ultralight vector field dark matter*, *JHEP* **02** (2017) 064, [arXiv:1611.03793].

[111] C. Armendariz-Picon, *Could dark energy be vector-like?*, *JCAP* **0407** (2004) 007, [astro-ph/0405267].

[112] J. Beltrn Jimnez and L. Heisenberg, *Non-trivial gravitational waves and structure formation phenomenology from dark energy*, *JCAP* **1809** (2018), no. 09 035, [arXiv:1806.01753].

[113] A. Golovnev, V. Mukhanov, and V. Vanchurin, *Vector Inflation*, *JCAP* **0806** (2008) 009, [arXiv:0802.2068].

[114] A. Proca, *Sur la theorie ondulatoire des electrons positifs et negatifs*, *J. Phys. Radium* **7** (1936) 347–353.

[115] G. Tasinato, *Cosmic Acceleration from Abelian Symmetry Breaking*, *JHEP* **04** (2014) 067, [arXiv:1402.6450].

[116] L. Heisenberg, *Generalization of the Proca Action*, *JCAP* **1405** (2014) 015, [arXiv:1402.7026].

[117] E. Allys, P. Peter, and Y. Rodriguez, *Generalized Proca action for an Abelian vector field*, *JCAP* **1602** (2016), no. 02 004, [arXiv:1511.03101].

[118] J. Beltran Jimenez and L. Heisenberg, *Derivative self-interactions for a massive vector field*, *Phys. Lett.* **B757** (2016) 405–411, [arXiv:1602.03410].

[119] L. Heisenberg, R. Kase, and S. Tsujikawa, *Beyond generalized Proca theories*, *Phys. Lett.* **B760** (2016) 617–626, [arXiv:1605.05565].

[120] R. Kimura, A. Naruko, and D. Yoshida, *Extended vector-tensor theories*, *JCAP* **1701** (2017), no. 01 002, [arXiv:1608.07066].

[121] E. Allys, P. Peter, and Y. Rodriguez, *Generalized SU(2) Proca Theory*, *Phys. Rev.* **D94** (2016), no. 8 084041, [arXiv:1609.05870].

[122] J. Beltran Jimenez and L. Heisenberg, *Generalized multi-Proca fields*, *Phys. Lett.* **B770** (2017) 16–26, [arXiv:1610.08960].

[123] L. Heisenberg, *Scalar-Vector-Tensor Gravity Theories*, *JCAP* **1810** (2018), no. 10 054, [arXiv:1801.01523].

[124] J. D. Bekenstein, *Relativistic gravitation theory for the MOND paradigm*, *Phys. Rev.* **D70** (2004) 083509, [astro-ph/0403694]. [Erratum: Phys. Rev.D71,069901(2005)].

[125] J.-P. Bruneton and G. Esposito-Farese, *Field-theoretical formulations of MOND-like gravity*, *Phys. Rev.* **D76** (2007) 124012, [arXiv:0705.4043]. [Erratum: Phys. Rev.D76,129902(2007)].

[126] E. Sagi, *Propagation of gravitational waves in generalized TeVeS*, *Phys. Rev.* **D81** (2010) 064031, [arXiv:1001.1555].

[127] C. Skordis, D. F. Mota, P. G. Ferreira, and C. Boehm, *Large Scale Structure in Bekenstein's theory of relativistic Modified Newtonian Dynamics*, *Phys. Rev. Lett.* **96** (2006) 011301, [astro-ph/0505519].

[128] F. Bourliot, P. G. Ferreira, D. F. Mota, and C. Skordis, *The cosmological behavior of Bekenstein's modified theory of gravity*, *Phys. Rev.* **D75** (2007) 063508, [astro-ph/0611255].

[129] C. Skordis, *The Tensor-Vector-Scalar theory and its cosmology*, *Class. Quant. Grav.* **26** (2009) 143001, [arXiv:0903.3602].

[130] C. de Rham, J. T. Deskins, A. J. Tolley, and S.-Y. Zhou, *Graviton Mass Bounds*, *Rev. Mod. Phys.* **89** (2017), no. 2 025004, [arXiv:1606.08462].

[131] M. Fierz and W. Pauli, *On relativistic wave equations for particles of arbitrary spin in an electromagnetic field*, *Proc. Roy. Soc. Lond.* **A173** (1939) 211–232.

[132] H. van Dam and M. J. G. Veltman, *Massive and massless Yang-Mills and gravitational fields*, *Nucl. Phys.* **B22** (1970) 397–411.

[133] V. I. Zakharov, *Linearized gravitation theory and the graviton mass*, *JETP Lett.* **12** (1970) 312. [Pisma Zh. Eksp. Teor. Fiz.12,447(1970)].

[134] A. I. Vainshtein, *To the problem of nonvanishing gravitation mass*, *Phys. Lett.* **B39** (1972) 393–394.

[135] D. G. Boulware and S. Deser, *Can gravitation have a finite range?*, *Phys. Rev.* **D6** (1972) 3368–3382.

[136] P. Creminelli, A. Nicolis, M. Papucci, and E. Trincherini, *Ghosts in massive gravity*, *JHEP* **09** (2005) 003, [hep-th/0505147].

[137] C. de Rham, G. Gabadadze, and A. J. Tolley, *Resummation of Massive Gravity*, *Phys. Rev. Lett.* **106** (2011) 231101, [arXiv:1011.1232].

[138] K. Hinterbichler, *Theoretical Aspects of Massive Gravity*, *Rev. Mod. Phys.* **84** (2012) 671–710, [arXiv:1105.3735].

[139] C. de Rham, *Massive Gravity*, *Living Rev. Rel.* **17** (2014) 7, [arXiv:1401.4173].

[140] S. F. Hassan and R. A. Rosen, *Resolving the Ghost Problem in non-Linear Massive Gravity*, *Phys. Rev. Lett.* **108** (2012) 041101, [arXiv:1106.3344].

[141] S. F. Hassan and R. A. Rosen, *Confirmation of the Secondary Constraint and Absence of Ghost in Massive Gravity and Bimetric Gravity*, *JHEP* **04** (2012) 123, [arXiv:1111.2070].

[142] E. Babichev and C. Deffayet, *An introduction to the Vainshtein mechanism*, *Class. Quant. Grav.* **30** (2013) 184001, [arXiv:1304.7240].

[143] E. Babichev and R. Brito, *Black holes in massive gravity*, *Class. Quant. Grav.* **32** (2015) 154001, [arXiv:1503.07529].

[144] G. D'Amico, C. de Rham, S. Dubovsky, G. Gabadadze, D. Pirtskhalava, and A. J. Tolley, *Massive Cosmologies*, *Phys. Rev.* **D84** (2011) 124046, [arXiv:1108.5231].

[145] A. E. Gumrukcuoglu, C. Lin, and S. Mukohyama, *Open FRW universes and self-acceleration from nonlinear massive gravity*, *JCAP* **1111** (2011) 030, [arXiv:1109.3845].

[146] A. De Felice, A. E. Gümrükçüoğlu, C. Lin, and S. Mukohyama, *Nonlinear stability of cosmological solutions in massive gravity*, *JCAP* **1305** (2013) 035, [arXiv:1303.4154].

[147] F. Könnig, H. Nersisyan, Y. Akrami, L. Amendola, and M. Zumalacárregui, *A spectre is haunting the cosmos: Quantum stability of massive gravity with ghosts*, *JHEP* **11** (2016) 118, [arXiv:1605.08757].

[148] C. de Rham, M. Fasiello, and A. J. Tolley, *Stable FLRW solutions in Generalized Massive Gravity*, *Int. J. Mod. Phys.* **D23** (2014), no. 13 1443006, [arXiv:1410.0960].

[149] G. D'Amico, G. Gabadadze, L. Hui, and D. Pirtskhalava, *Quasidilaton: Theory and cosmology*, *Phys. Rev.* **D87** (2013) 064037, [arXiv:1206.4253].

[150] S. F. Hassan and R. A. Rosen, *Bimetric Gravity from Ghost-free Massive Gravity*, *JHEP* **02** (2012) 126, [arXiv:1109.3515].

[151] K. Hinterbichler and R. A. Rosen, *Interacting Spin-2 Fields*, *JHEP* **07** (2012) 047, [arXiv:1203.5783].

[152] Y. Akrami, T. S. Koivisto, and M. Sandstad, *Accelerated expansion from ghost-free bigravity: a statistical analysis with improved generality*, *JHEP* **03** (2013) 099, [arXiv:1209.0457].

[153] M. Fasiello and A. J. Tolley, *Cosmological Stability Bound in Massive Gravity and Bigravity*, *JCAP* **1312** (2013) 002, [arXiv:1308.1647].

[154] D. Comelli, M. Crisostomi, and L. Pilo, *FRW Cosmological Perturbations in Massive Bigravity*, *Phys. Rev.* **D90** (2014) 084003, [arXiv:1403.5679].

[155] F. Könnig, *Higuchi Ghosts and Gradient Instabilities in Bimetric Gravity*, *Phys. Rev.* **D91** (2015) 104019, [arXiv:1503.07436].

[156] E. Mortsell and J. Enander, *Scalar instabilities in bimetric gravity: The Vainshtein mechanism and structure formation*, *JCAP* **1510** (2015), no. 10 044, [arXiv:1506.04977].

[157] Y. Akrami, S. F. Hassan, F. Könnig, A. Schmidt-May, and A. R. Solomon, *Bimetric gravity is cosmologically viable*, *Phys. Lett.* **B748** (2015) 37–44, [arXiv:1503.07521].

[158] C.-P. Ma and E. Bertschinger, *Cosmological perturbation theory in the synchronous and conformal Newtonian gauges*, *Astrophys. J.* **455** (1995) 7–25, [astro-ph/9506072].

[159] M. Chevallier and D. Polarski, *Accelerating universes with scaling dark matter*, *Int. J. Mod. Phys.* **D10** (2001) 213–224, [gr-qc/0009008].

[160] E. V. Linder, *Exploring the expansion history of the universe*, *Phys. Rev. Lett.* **90** (2003) 091301, [astro-ph/0208512].

[161] E. J. Ruiz and D. Huterer, *Testing the dark energy consistency with geometry and growth*, *Phys. Rev.* **D91** (2015) 063009, [arXiv:1410.5832].

[162] J. L. Bernal, L. Verde, and A. J. Cuesta, *Parameter splitting in dark energy: is dark energy the same in the background and in the cosmic structures?*, *JCAP* **1602** (2016), no. 02 059, [arXiv:1511.03049].

[163] G. Gubitosi, F. Piazza, and F. Vernizzi, *The Effective Field Theory of Dark Energy*, *JCAP* **1302** (2013) 032, [arXiv:1210.0201]. [JCAP1302,032(2013)].

[164] J. K. Bloomfield, É. É. Flanagan, M. Park, and S. Watson, *Dark energy or modified gravity? An effective field theory approach*, *JCAP* **1308** (2013) 010, [arXiv:1211.7054].

[165] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, *Essential Building Blocks of Dark Energy*, *JCAP* **1308** (2013) 025, [arXiv:1304.4840].

[166] J. Gleyzes, D. Langlois, and F. Vernizzi, *A unifying description of dark energy*, *Int. J. Mod. Phys.* **D23** (2015), no. 13 1443010, [arXiv:1411.3712].

[167] E. Bellini and I. Sawicki, *Maximal freedom at minimum cost: linear large-scale structure in general modifications of gravity*, *JCAP* **1407** (2014) 050, [arXiv:1404.3713].

[168] D. Bettoni and M. Zumalacárregui, *Kinetic mixing in scalar-tensor theories of gravity*, *Phys. Rev.* **D91** (2015) 104009, [arXiv:1502.02666].

[169] J. Kennedy, L. Lombriser, and A. Taylor, *Reconstructing Horndeski models from the effective field theory of dark energy*, *Phys. Rev.* **D96** (2017), no. 8 084051, [arXiv:1705.09290].

[170] J. Kennedy, L. Lombriser, and A. Taylor, *Reconstructing Horndeski theories from phenomenological modified gravity and dark energy models on cosmological scales*, *Phys. Rev.* **D98** (2018), no. 4 044051, [arXiv:1804.04582].

[171] G. D'Amico, Z. Huang, M. Mancarella, and F. Vernizzi, *Weakening Gravity on Redshift-Survey Scales with Kinetic Matter Mixing*, *JCAP* **1702** (2017) 014, [arXiv:1609.01272].

[172] M. Lagos, T. Baker, P. G. Ferreira, and J. Noller, *A general theory of linear cosmological perturbations: scalar-tensor and vector-tensor theories*, *JCAP* **1608** (2016), no. 08 007, [arXiv:1604.01396].

[173] M. Lagos and P. G. Ferreira, *A general theory of linear cosmological perturbations: bimetric theories*, *JCAP* **1701** (2017), no. 01 047, [arXiv:1610.00553].

[174] M. Lagos, E. Bellini, J. Noller, P. G. Ferreira, and T. Baker, *A general theory of linear cosmological perturbations: stability conditions, the quasistatic limit and dynamics*, *JCAP* **1803** (2018), no. 03 021, [arXiv:1711.09893].

[175] B. Hu, M. Raveri, N. Frusciante, and A. Silvestri, *Effective Field Theory of Cosmic Acceleration: an implementation in CAMB*, *Phys. Rev.* **D89** (2014), no. 10 103530, [arXiv:1312.5742].

[176] M. Zumalacárregui, E. Bellini, I. Sawicki, J. Lesgourgues, and P. G. Ferreira, *hiclass: Horndeski in the Cosmic Linear Anisotropy Solving System*, *JCAP* **1708** (2017), no. 08 019, [arXiv:1605.06102].

[177] Z. Huang, *Observational effects of a running Planck mass*, *Phys. Rev.* **D93** (2016), no. 4 043538, [arXiv:1511.02808].

[178] A. Lewis, A. Challinor, and A. Lasenby, *Efficient computation of CMB anisotropies in closed FRW models*, *Astrophys. J.* **538** (2000) 473–476, [astro-ph/9911177].

[179] D. Blas, J. Lesgourgues, and T. Tram, *The Cosmic Linear Anisotropy Solving System (CLASS) II: Approximation schemes*, *JCAP* **1107** (2011) 034, [arXiv:1104.2933].

[180] E. Di Dio, F. Montanari, J. Lesgourgues, and R. Durrer, *The CLASSgal code for Relativistic Cosmological Large Scale Structure*, *JCAP* **1311** (2013) 044, [arXiv:1307.1459].

[181] E. Bellini et al., *Comparison of Einstein-Boltzmann solvers for testing general relativity*, *Phys. Rev.* **D97** (2018), no. 2 023520, [arXiv:1709.09135].

[182] E. Bellini, A. J. Cuesta, R. Jimenez, and L. Verde, *Constraints on deviations from LambdaCDM within Horndeski gravity*, *JCAP* **1602** (2016), no. 02 053, [arXiv:1509.07816]. [Erratum: JCAP1606,no.06,E01(2016)].

[183] **Planck** Collaboration, P. A. R. Ade et al., *Planck 2015 results. XIV. Dark energy and modified gravity*, *Astron. Astrophys.* **594** (2016) A14, [arXiv:1502.01590].

[184] C. D. Kreisch and E. Komatsu, *Cosmological Constraints on Horndeski Gravity in Light of GW170817*, *JCAP* **1812** (2018), no. 12 030, [arXiv:1712.02710].

[185] J. Gleyzes, D. Langlois, M. Mancarella, and F. Vernizzi, *Effective Theory of Dark Energy at Redshift Survey Scales*, *JCAP* **1602** (2016), no. 02 056, [arXiv:1509.02191].

[186] D. Alonso, E. Bellini, P. G. Ferreira, and M. Zumalacárregui, *Observational future of cosmological scalar-tensor theories*, *Phys. Rev.* **D95** (2017), no. 6 063502, [arXiv:1610.09290].

[187] C. S. Lorenz, D. Alonso, and P. G. Ferreira, *Impact of relativistic effects on cosmological parameter estimation*, *Phys. Rev.* **D97** (2018), no. 2 023537, [arXiv:1710.02477].

[188] A. Spurio Mancini, R. Reischke, V. Pettorino, B. M. Schfer, and M. Zumalacrregui, *Testing (modified) gravity with 3D and tomographic cosmic shear*, *Mon. Not. Roy. Astron. Soc.* **480** (2018) 3725, [arXiv:1801.04251].

[189] R. Reischke, A. Spurio Mancini, B. M. Schfer, and P. M. Merkel, *Investigating scalartensor gravity with statistics of the cosmic large-scale structure*, *Mon. Not. Roy. Astron. Soc.* **482** (2019), no. 3 3274–3287, [arXiv:1804.02441].

[190] J. Renk, M. Zumalacarregui, and F. Montanari, *Gravity at the horizon: on relativistic effects, CMB-LSS correlations and ultra-large scales in Horndeski's theory*, *JCAP* **1607** (2016), no. 07 040, [arXiv:1604.03487].

[191] J. Gleyzes, *Parametrizing modified gravity for cosmological surveys*, *Phys. Rev.* **D96** (2017), no. 6 063516, [arXiv:1705.04714].

[192] E. Bellini, R. Jimenez, and L. Verde, *Signatures of Horndeski gravity on the Dark Matter Bispectrum*, *JCAP* **1505** (2015), no. 05 057, [arXiv:1504.04341].

[193] G. Cusin, M. Lewandowski, and F. Vernizzi, *Nonlinear Effective Theory of Dark Energy*, *JCAP* **1804** (2018), no. 04 061, [arXiv:1712.02782].

[194] G. Cusin, M. Lewandowski, and F. Vernizzi, *Dark Energy and Modified Gravity in the Effective Field Theory of Large-Scale Structure*, *JCAP* **1804** (2018), no. 04 005, [arXiv:1712.02783].

[195] D. Yamauchi, S. Yokoyama, and H. Tashiro, *Constraining modified theories of gravity with the galaxy bispectrum*, *Phys. Rev.* **D96** (2017), no. 12 123516, [arXiv:1709.03243].

[196] T. Namikawa, F. R. Bouchet, and A. Taruya, *CMB lensing bispectrum as a probe of modified gravity theories*, *Phys. Rev.* **D98** (2018), no. 4 043530, [arXiv:1805.10567].

[197] E. Bellini and M. Zumalacarregui, *Nonlinear evolution of the baryon acoustic oscillation scale in alternative theories of gravity*, *Phys. Rev.* **D92** (2015), no. 6 063522, [arXiv:1505.03839].

[198] A. De Felice, T. Kobayashi, and S. Tsujikawa, *Effective gravitational couplings for cosmological perturbations in the most general scalar-tensor theories with second-order field equations*, *Phys. Lett.* **B706** (2011) 123–133, [arXiv:1108.4242].

[199] L. Amendola, M. Kunz, M. Motta, I. D. Saltas, and I. Sawicki, *Observables and unobservables in dark energy cosmologies*, *Phys. Rev.* **D87** (2013), no. 2 023501, [arXiv:1210.0439].

[200] I. Sawicki and E. Bellini, *Limits of quasistatic approximation in modified-gravity cosmologies*, *Phys. Rev.* **D92** (2015), no. 8 084061, [arXiv:1503.06831].

[201] A. Silvestri, L. Pogosian, and R. V. Buniy, *Practical approach to cosmological perturbations in modified gravity*, *Phys. Rev.* **D87** (2013), no. 10 104015, [arXiv:1302.1193].

[202] T. Baker, P. G. Ferreira, C. D. Leonard, and M. Motta, *New Gravitational Scales in Cosmological Surveys*, *Phys. Rev.* **D90** (2014), no. 12 124030, [arXiv:1409.8284].

[203] V. Vardanyan and L. Amendola, *How can we tell whether dark energy is composed of multiple fields?*, *Phys. Rev.* **D92** (2015), no. 2 024009, [arXiv:1502.05922].

[204] L. Amendola, S. Fogli, A. Guarnizo, M. Kunz, and A. Vollmer, *Model-independent constraints on the cosmological anisotropic stress*, *Phys. Rev.* **D89** (2014), no. 6 063538, [arXiv:1311.4765].

[205] L. Taddei, M. Martinelli, and L. Amendola, *Model-independent constraints on modified gravity from current data and from the Euclid and SKA future surveys*, *JCAP* **1612** (2016), no. 12 032, [arXiv:1604.01059].

[206] T. Baker and P. Bull, *Observational signatures of modified gravity on ultra-large scales*, *Astrophys. J.* **811** (2015) 116, [arXiv:1506.00641].

[207] E. Villa, E. Di Dio, and F. Lepori, *Lensing convergence in galaxy clustering in LambdaCDM and beyond*, *JCAP* **1804** (2018), no. 04 033, [arXiv:1711.07466].

[208] A. Raccanelli, D. Bertacca, O. Doré, and R. Maartens, *Large-scale 3D galaxy correlation function and non-Gaussianity*, *JCAP* **1408** (2014) 022, [arXiv:1306.6646].

[209] L. Lombriser, J. Yoo, and K. Koyama, *Relativistic effects in galaxy clustering in a parametrized post-Friedmann universe*, *Phys. Rev.* **D87** (2013) 104019, [arXiv:1301.3132].

[210] C. Bonvin and P. Fleury, *Testing the equivalence principle on cosmological scales*, *JCAP* **1805** (2018), no. 05 061, [arXiv:1803.02771].

[211] I. D. Saltas, I. Sawicki, L. Amendola, and M. Kunz, *Anisotropic Stress as a Signature of Nonstandard Propagation of Gravitational Waves*, *Phys. Rev. Lett.* **113** (2014), no. 19 191101, [arXiv:1406.7139].

[212] L. Perenon, F. Piazza, C. Marinoni, and L. Hui, *Phenomenology of dark energy: general features of large-scale perturbations*, *JCAP* **1511** (2015), no. 11 029, [arXiv:1506.03047].

[213] S. Peirone, K. Koyama, L. Pogosian, M. Raveri, and A. Silvestri, *Large-scale structure phenomenology of viable Horndeski theories*, *Phys. Rev.* **D97** (2018), no. 4 043519, [arXiv:1712.00444].

[214] L. Pogosian and A. Silvestri, *What can cosmology tell us about gravity? Constraining Horndeski gravity with Σ and μ*, *Phys. Rev.* **D94** (2016), no. 10 104014, [arXiv:1606.05339].

[215] D. J. E. Marsh, P. Bull, P. G. Ferreira, and A. Pontzen, *Quintessence in a quandary: Prior dependence in dark energy models*, *Phys. Rev.* **D90** (2014), no. 10 105023, [arXiv:1406.2301].

[216] M. Raveri, P. Bull, A. Silvestri, and L. Pogosian, *Priors on the effective Dark Energy equation of state in scalar-tensor theories*, *Phys. Rev.* **D96** (2017), no. 8 083509, [arXiv:1703.05297].

[217] P. Peter and J.-P. Uzan, *Primordial cosmology*. Oxford University Press, 2013.

[218] J. Silk, *Challenges in Cosmology from the Big Bang to Dark Energy, Dark Matter and Galaxy Formation*, *JPS Conf. Proc.* **14** (2017) 010101, [arXiv:1611.09846].

[219] J. Kormendy and D. Richstone, *Inward bound: The Search for supermassive black holes in galactic nuclei*, *Ann. Rev. Astron. Astrophys.* **33** (1995) 581.

[220] G. Bertone, *Particle dark matter: observations, models and searches*. Cambridge University Press, 2010.

[221] L. Hui, J. P. Ostriker, S. Tremaine, and E. Witten, *Ultralight scalars as cosmological dark matter*, *Phys. Rev.* **D95** (2017), no. 4 043541, [arXiv:1610.08297].

[222] S. Hawking, *Gravitationally collapsed objects of very low mass*, *Mon. Not. Roy. Astron. Soc.* **152** (1971) 75.

[223] B. J. Carr and S. W. Hawking, *Black holes in the early Universe*, *Mon. Not. Roy. Astron. Soc.* **168** (1974) 399–415.

[224] S. Clesse and J. García-Bellido, *Massive Primordial Black Holes from Hybrid Inflation as Dark Matter and the seeds of Galaxies*, *Phys. Rev.* **D92** (2015), no. 2 023524, [arXiv:1501.07565].

[225] B. Carr and J. Silk, *Primordial Black Holes as Generators of Cosmic Structures*, *Mon. Not. Roy. Astron. Soc.* **478** (2018), no. 3 3756–3775, [arXiv:1801.00672].

[226] B. Carr, F. Kuhnel, and M. Sandstad, *Primordial Black Holes as Dark Matter*, *Phys. Rev.* **D94** (2016), no. 8 083504, [arXiv:1607.06077].

[227] B. Paczynski, *Gravitational microlensing by the galactic halo*, *Astrophys. J.* **304** (1986) 1–5.

[228] **Macho** Collaboration, R. A. Allsman et al., *MACHO project limits on black hole dark matter in the 1-30 solar mass range*, *Astrophys. J.* **550** (2001) L169, [astro-ph/0011506].

[229] **EROS-2** Collaboration, P. Tisserand et al., *Limits on the Macho Content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds*, *Astron. Astrophys.* **469** (2007) 387–404, [astro-ph/0607207].

[230] L. Wyrzykowski et al., *The OGLE View of Microlensing towards the Magellanic Clouds. IV. OGLE-III SMC Data and Final Conclusions on MACHOs*, *Mon. Not. Roy. Astron. Soc.* **416** (2011) 2949, [arXiv:1106.2925].

[231] H. Niikura et al., *Microlensing constraints on primordial black holes with the Subaru/HSC Andromeda observation*, *Nat. Astron.* **3** (2019), no. 6 524–534, [arXiv:1701.02151].

[232] K. Griest, A. M. Cieplak, and M. J. Lehner, *Experimental Limits on Primordial Black Hole Dark Matter from the First 2 yr of Kepler Data*, *Astrophys. J.* **786** (2014), no. 2 158, [arXiv:1307.5798].

[233] A. Barnacka, J. F. Glicenstein, and R. Moderski, *New constraints on primordial black holes abundance from femtolensing of gamma-ray bursts*, *Phys. Rev.* **D86** (2012) 043001, [arXiv:1204.2056].

[234] A. Katz, J. Kopp, S. Sibiryakov, and W. Xue, *Femtolensing by Dark Matter Revisited*, *JCAP* **1812** (2018) 005, [arXiv:1807.11495].

[235] M. Zumalacarregui and U. Seljak, *Limits on stellar-mass compact objects as dark matter from gravitational lensing of type Ia supernovae*, *Phys. Rev. Lett.* **121** (2018), no. 14 141101, [arXiv:1712.02240].

[236] M. A. Monroy-Rodrguez and C. Allen, *The end of the MACHO era- revisited: new limits on MACHO masses from halo wide binaries*, *Astrophys. J.* **790** (2014), no. 2 159, [arXiv:1406.5169].

[237] T. D. Brandt, *Constraints on MACHO Dark Matter from Compact Stellar Systems in Ultra-Faint Dwarf Galaxies*, *Astrophys. J.* **824** (2016), no. 2 L31, [arXiv:1605.03665].

[238] T. Nakamura, M. Sasaki, T. Tanaka, and K. S. Thorne, *Gravitational waves from coalescing black hole MACHO binaries*, *Astrophys. J.* **487** (1997) L139–L142, [astro-ph/9708060].

[239] M. Sasaki, T. Suyama, T. Tanaka, and S. Yokoyama, *Primordial Black Hole Scenario for the Gravitational-Wave Event GW150914*, *Phys. Rev. Lett.* **117** (2016), no. 6 061101, [arXiv:1603.08338].

[240] S. Bird, I. Cholis, J. B. Muoz, Y. Ali-Hamoud, M. Kamionkowski, E. D. Kovetz, A. Raccanelli, and A. G. Riess, *Did LIGO detect dark matter?*, *Phys. Rev. Lett.* **116** (2016), no. 20 201301, [arXiv:1603.00464].

[241] S. Clesse and J. García-Bellido, *The clustering of massive Primordial Black Holes as Dark Matter: measuring their mass distribution with Advanced LIGO*, *Phys. Dark Univ.* **15** (2017) 142–147, [arXiv:1603.05234].

[242] B. Zackay, T. Venumadhav, L. Dai, J. Roulet, and M. Zaldarriaga, *A Highly Spinning and Aligned Binary Black Hole Merger in the Advanced LIGO First Observing Run*, arXiv:1902.10331.

[243] T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga, *New Binary Black Hole Mergers in the Second Observing Run of Advanced LIGO and Advanced Virgo*, arXiv:1904.07214.

[244] K. Belczynski, D. E. Holz, T. Bulik, and R. O'Shaughnessy, *The first gravitational-wave source from the isolated evolution of two 40-100 Msun stars*, *Nature* **534** (2016) 512, [arXiv:1602.04531].

[245] K. Belczynski et al., *The origin of low spin of black holes in LIGO/Virgo mergers*, arXiv:1706.07053.

[246] T. Chiba and S. Yokoyama, *Spin Distribution of Primordial Black Holes*, *PTEP* **2017** (2017), no. 8 083E01, [arXiv:1704.06573].

[247] Y. Ali-Hamoud and M. Kamionkowski, *Cosmic microwave background limits on accreting primordial black holes*, *Phys. Rev.* **D95** (2017), no. 4 043534, [arXiv:1612.05644].

[248] B. J. Carr, K. Kohri, Y. Sendouda, and J. Yokoyama, *New cosmological constraints on primordial black holes*, *Phys. Rev.* **D81** (2010) 104019, [arXiv:0912.5297].

[249] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*. Macmillan, 1973.

[250] M. Maggiore, *Gravitational Waves: Volume 1: Theory and Experiments*, vol. 1. Oxford university press, 2008.

[251] M. Maggiore, *Gravitational Waves. Vol. 2: Astrophysics and Cosmology*. Oxford University Press, 2018.

[252] E. E. Flanagan and S. A. Hughes, *The Basics of gravitational wave theory*, *New J. Phys.* **7** (2005) 204, [gr-qc/0501041].

[253] S. M. Carroll, *Spacetime and geometry: An introduction to general relativity*. Addison-Wesley, 2004.

[254] M. Vallisneri, J. Kanner, R. Williams, A. Weinstein, and B. Stephens, *The LIGO Open Science Center*, *J. Phys. Conf. Ser.* **610** (2015), no. 1 012021, [arXiv:1410.4839].

[255] C. Caprini and D. G. Figueroa, *Cosmological Backgrounds of Gravitational Waves*, *Class. Quant. Grav.* **35** (2018), no. 16 163001, [arXiv:1801.04268].

[256] G. Baym, S. P. Patil, and C. J. Pethick, *Damping of gravitational waves by matter*, *Phys. Rev.* **D96** (2017), no. 8 084033, [arXiv:1707.05192].

[257] R. Flauger and S. Weinberg, *Gravitational Waves in Cold Dark Matter*, *Phys. Rev.* **D97** (2018), no. 12 123506, [arXiv:1801.00386].

[258] W. L. Freedman, *Cosmology at a Crossroads*, *Nat. Astron.* **1** (2017) 0121, [arXiv:1706.02739].

[259] R. L. Beaton et al., *The Carnegie-Chicago Hubble Program. I. An Independent Approach to the Extragalactic Distance Scale Using only Population II Distance Indicators*, *Astrophys. J.* **832** (2016), no. 2 210, [arXiv:1604.01788].

[260] A. G. Riess et al., *Milky Way Cepheid Standards for Measuring Cosmic Distances and Application to Gaia DR2: Implications for the Hubble Constant*, *Astrophys. J.* **861** (2018), no. 2 126, [arXiv:1804.10655].

[261] **CMB-S4** Collaboration, K. N. Abazajian et al., *CMB-S4 Science Book, First Edition*, arXiv:1610.02743.

[262] S. Nissanke, D. E. Holz, N. Dalal, S. A. Hughes, J. L. Sievers, and C. M. Hirata, *Determining the Hubble constant from gravitational wave observations of merging compact binaries*, arXiv:1307.2638.

[263] S. Casertano et al., *Parallax of Galactic Cepheids from Spatially Scanning the Wide Field Camera 3 on the Hubble Space Telescope: The Case of SS Canis Majoris*, *Astrophys. J.* **825** (2016), no. 1 11, [arXiv:1512.09371].

[264] A. G. Riess et al., *A 2.4% Determination of the Local Value of the Hubble Constant*, *Astrophys. J.* **826** (2016), no. 1 56, [arXiv:1604.01424].

[265] B. F. Schutz, *Determining the Hubble Constant from Gravitational Wave Observations*, *Nature* **323** (1986) 310–311.

[266] R. Perna, D. Lazzati, and B. Giacomazzo, *Short Gamma-Ray Bursts from the Merger of Two Black Holes*, *Astrophys. J.* **821** (2016), no. 1 L18, [arXiv:1602.05140].

[267] **Virgo, LIGO Scientific** Collaboration, B. P. Abbott et al., *Properties of the Binary Black Hole Merger GW150914*, *Phys. Rev. Lett.* **116** (2016), no. 24 241102, [arXiv:1602.03840].

[268] B. D. Metzger, *Kilonovae*, *Living Rev. Rel.* **20** (2017) 3, [arXiv:1610.09381].

[269] N. Dalal, D. E. Holz, S. A. Hughes, and B. Jain, *Short grb and binary black hole standard sirens as a probe of dark energy*, *Phys. Rev.* **D74** (2006) 063006, [astro-ph/0601275].

[270] S. Vitale and H.-Y. Chen, *Measuring the Hubble constant with neutron star black hole mergers*, *Phys. Rev. Lett.* **121** (2018), no. 2 021303, [arXiv:1804.07337].

[271] D. E. Holz and S. A. Hughes, *Using gravitational-wave standard sirens*, *Astrophys. J.* **629** (2005) 15–22, [astro-ph/0504616].

[272] S. Nissanke, D. E. Holz, S. A. Hughes, N. Dalal, and J. L. Sievers, *Exploring short gamma-ray bursts as gravitational-wave standard sirens*, *Astrophys. J.* **725** (2010) 496–514, [arXiv:0904.1017].

[273] B. S. Sathyaprakash, B. F. Schutz, and C. Van Den Broeck, *Cosmography with the Einstein Telescope*, *Class. Quant. Grav.* **27** (2010) 215006, [arXiv:0906.4151].

[274] W. Del Pozzo, *Inference of the cosmological parameters from gravitational waves: application to second generation interferometers*, *Phys. Rev.* **D86** (2012) 043011, [arXiv:1108.1317].

[275] H.-Y. Chen and D. E. Holz, *Finding the One: Identifying the Host Galaxies of Gravitational-Wave Sources*, arXiv:1612.01471.

[276] C. Messenger and J. Read, *Measuring a cosmological distance-redshift relationship using only gravitational wave observations of binary neutron star coalescences*, *Phys. Rev. Lett.* **108** (2012) 091101, [arXiv:1107.5725].

[277] W. Del Pozzo, T. G. F. Li, and C. Messenger, *Cosmological inference using only gravitational wave observations of binary neutron stars*, *Phys. Rev.* **D95** (2017), no. 4 043502, [arXiv:1506.06590].

[278] S. R. Taylor, J. R. Gair, and I. Mandel, *Hubble without the Hubble: Cosmology using advanced gravitational-wave detectors alone*, *Phys. Rev.* **D85** (2012) 023535, [arXiv:1108.5161].

[279] C. Messenger, K. Takami, S. Gossan, L. Rezzolla, and B. S. Sathyaprakash, *Source Redshifts from Gravitational-Wave Observations of Binary Neutron Star Mergers*, *Phys. Rev.* **X4** (2014), no. 4 041004, [arXiv:1312.1862].

[280] C. Guidorzi et al., *Improved Constraints on H_0 from a Combined Analysis of Gravitational-wave and Electromagnetic Emission from GW170817*, *Astrophys. J.* **851** (2017), no. 2 L36, [arXiv:1710.06426].

[281] K. Hotokezaka, E. Nakar, O. Gottlieb, S. Nissanke, K. Masuda, G. Hallinan, K. P. Mooley, and A. Deller, *A Hubble constant measurement from superluminal motion of the jet in GW170817*, arXiv:1806.10596.

[282] **Virgo, members of the LIGO Scientific** Collaboration, M. Fishbach, R. Gray, I. M. Hernandez, H. Qi, and A. Sur, *A standard siren measurement of the Hubble constant from GW170817 without the electromagnetic counterpart*, arXiv:1807.05667.

[283] H.-Y. Chen, M. Fishbach, and D. E. Holz, *A two per cent Hubble constant measurement from standard sirens within five years*, *Nature* **562** (2018), no. 7728 545–547, [arXiv:1712.06531].

[284] S. M. Feeney, H. V. Peiris, A. R. Williamson, S. M. Nissanke, D. J. Mortlock, J. Alsing, and D. Scolnic, *Prospects for resolving the Hubble constant tension with standard sirens*, *Phys. Rev. Lett.* **122** (2019), no. 6 061105, [arXiv:1802.03404].

[285] N. Tamanini, C. Caprini, E. Barausse, A. Sesana, A. Klein, and A. Petiteau, *Science with the space-based interferometer eLISA. III: Probing the expansion of the Universe using gravitational wave standard sirens*, *JCAP* **1604** (2016), no. 04 002, [arXiv:1601.07112].

[286] W. Cardona, M. Kunz, and V. Pettorino, *Determining H_0 with Bayesian hyper-parameters*, *JCAP* **1703** (2017), no. 03 056, [arXiv:1611.06088].

[287] S. M. Feeney, D. J. Mortlock, and N. Dalmasso, *Clarifying the Hubble constant tension with a Bayesian hierarchical model of the local distance ladder*, *Mon. Not. Roy. Astron. Soc.* **476** (2018), no. 3 3861–3882, [arXiv:1707.00007].

[288] V. Bonvin et al., *H0LiCOW - V. New COSMOGRAIL time delays of HE 0435-1223: H0 to 3.8 per cent precision from strong lensing in a flat LCDM model*, *Mon. Not. Roy. Astron. Soc.* **465** (2017), no. 4 4914–4930, [arXiv:1607.01790].

[289] G. E. Addison, D. J. Watts, C. L. Bennett, M. Halpern, G. Hinshaw, and J. L. Weiland, *Elucidating ΛCDM: Impact of Baryon Acoustic Oscillation Measurements on the Hubble Constant Discrepancy*, *Astrophys. J.* **853** (2018), no. 2 119, [arXiv:1707.06547].

[290] S. H. Suyu, T.-C. Chang, F. Courbin, and T. Okumura, *Cosmological distance indicators*, *Space Sci. Rev.* **214** (2018), no. 5 91, [arXiv:1801.07262].

[291] J. L. Bernal and J. A. Peacock, *Conservative cosmology: combining data with allowance for unknown systematics*, *JCAP* **1807** (2018), no. 07 002, [arXiv:1803.04470].

[292] A. J. Cuesta, L. Verde, A. Riess, and R. Jimenez, *Calibrating the cosmic distance scale ladder: the role of the sound horizon scale and the local expansion rate as distance anchors*, *Mon. Not. Roy. Astron. Soc.* **448** (2015), no. 4 3463–3471, [arXiv:1411.1094].

[293] J. L. Bernal, L. Verde, and A. G. Riess, *The trouble with H_0*, *JCAP* **1610** (2016), no. 10 019, [arXiv:1607.05617].

[294] V. Poulin, K. K. Boddy, S. Bird, and M. Kamionkowski, *Implications of an extended dark energy cosmology with massive neutrinos for cosmological tensions*, *Phys. Rev.* **D97** (2018), no. 12 123504, [arXiv:1803.02474].

[295] E. Di Valentino, A. Melchiorri, and O. Mena, *Can interacting dark energy solve the H_0 tension?*, *Phys. Rev.* **D96** (2017), no. 4 043503, [arXiv:1704.08342].

[296] A. De Felice and S. Tsujikawa, *Cosmology of a covariant Galileon field*, *Phys. Rev. Lett.* **105** (2010) 111301, [arXiv:1007.2700].

[297] A. Barreira, B. Li, A. Sanchez, C. M. Baugh, and S. Pascoli, *Parameter space in Galileon gravity models*, *Phys. Rev.* **D87** (2013) 103511, [arXiv:1302.6241].

[298] A. Barreira, B. Li, C. Baugh, and S. Pascoli, *The observational status of Galileon gravity after Planck*, *JCAP* **1408** (2014) 059, [arXiv:1406.0485].

[299] J. Renk, M. Zumalacárregui, F. Montanari, and A. Barreira, *Galileon gravity in light of ISW, CMB, BAO and H_0 data*, *JCAP* **1710** (2017), no. 10 020, [arXiv:1707.02263].

[300] S. Peirone, N. Frusciante, B. Hu, M. Raveri, and A. Silvestri, *Do current cosmological observations rule out all Covariant Galileons?*, *Phys. Rev.* **D97** (2018), no. 6 063518, [arXiv:1711.04760].

[301] E. Belgacem, Y. Dirian, S. Foffa, and M. Maggiore, *Nonlocal gravity. Conceptual aspects and cosmological predictions*, *JCAP* **1803** (2018), no. 03 002, [arXiv:1712.07066].

[302] R. A. Hulse and J. H. Taylor, *Discovery of a pulsar in a binary system*, *Astrophys. J.* **195** (1975) L51–L53.

[303] J. M. Weisberg, D. J. Nice, and J. H. Taylor, *Timing Measurements of the Relativistic Binary Pulsar PSR B1913+16*, *Astrophys. J.* **722** (2010) 1030–1034, [arXiv:1011.0718].

[304] I. H. Stairs, *Testing general relativity with pulsar timing*, *Living Rev. Rel.* **6** (2003) 5, [astro-ph/0307536].

[305] N. Wex, *Testing Relativistic Gravity with Radio Pulsars*, arXiv:1402.5594.

[306] T. Jacobson and D. Mattingly, *Einstein-Aether waves*, *Phys. Rev.* **D70** (2004) 024003, [gr-qc/0402005].

[307] K. Yagi, D. Blas, N. Yunes, and E. Barausse, *Strong Binary Pulsar Constraints on Lorentz Violation in Gravity*, *Phys. Rev. Lett.* **112** (2014), no. 16 161101, [arXiv:1307.6219].

[308] K. Yagi, D. Blas, E. Barausse, and N. Yunes, *Constraints on Einstein-Æther theory and Hořava gravity from binary pulsar observations*, *Phys. Rev.* **D89** (2014), no. 8 084067, [arXiv:1311.7144]. [Erratum: Phys. Rev.D90,no.6,069901(2014)].

[309] P. C. C. Freire, N. Wex, G. Esposito-Farese, J. P. W. Verbiest, M. Bailes, B. A. Jacoby, M. Kramer, I. H. Stairs, J. Antoniadis, and G. H. Janssen, *The relativistic pulsar-white dwarf binary PSR J1738+0333 II. The most stringent test of scalar-tensor gravity*, *Mon. Not. Roy. Astron. Soc.* **423** (2012) 3328, [arXiv:1205.1450].

[310] J. A. R. Cembranos, M. Coma Daz, and P. Martn-Moruno, *Modified gravity as a diagravitational medium*, *Phys. Lett.* **B788** (2019) 336–340, [arXiv:1805.09629].

[311] **VIRGO, KAGRA, LIGO Scientific** Collaboration, B. P. Abbott et al., *Prospects for Observing and Localizing Gravitational-Wave Transients with Advanced LIGO, Advanced Virgo and KAGRA*, *Living Rev. Rel.* **21** (2018) 3, [arXiv:1304.0670]. [Living Rev. Rel.19,1(2016)].

[312] B. Sathyaprakash et al., *Scientific Objectives of Einstein Telescope*, *Class. Quant. Grav.* **29** (2012) 124013, [arXiv:1206.0331]. [Erratum: Class. Quant. Grav.30,079501(2013)].

[313] **LIGO Scientific** Collaboration, B. P. Abbott et al., *Exploring the Sensitivity of Next Generation Gravitational Wave Detectors*, *Class. Quant. Grav.* **34** (2017), no. 4 044001, [arXiv:1607.08697].

[314] P. Amaro-Seoane et al., *eLISA/NGO: Astrophysics and cosmology in the gravitational-wave millihertz regime*, *GW Notes* **6** (2013) 4–110, [arXiv:1201.3621].

[315] C. J. Moore, R. H. Cole, and C. P. L. Berry, *Gravitational-wave sensitivity curves*, *Class. Quant. Grav.* **32** (2015), no. 1 015014, [arXiv:1408.0740].

[316] T. Narikawa, K. Ueno, H. Tagoshi, T. Tanaka, N. Kanda, and T. Nakamura, *Detectability of bigravity with graviton oscillations using gravitational wave observations*, *Phys. Rev.* **D91** (2015) 062007, [arXiv:1412.8074].

[317] K. Max, M. Platscher, and J. Smirnov, *Gravitational Wave Oscillations in Bigravity*, *Phys. Rev. Lett.* **119** (2017), no. 11 111101, [arXiv:1703.07785].

[318] K. Max, M. Platscher, and J. Smirnov, *Decoherence of Gravitational Wave Oscillations in Bigravity*, *Phys. Rev.* **D97** (2018), no. 6 064009, [arXiv:1712.06601].

[319] R. R. Caldwell, C. Devulder, and N. A. Maksimova, *Gravitational wave–Gauge field oscillations*, *Phys. Rev.* **D94** (2016), no. 6 063005, [arXiv:1604.08939].

[320] **KAGRA** Collaboration, K. Somiya, *Detector configuration of KAGRA: The Japanese cryogenic gravitational-wave detector*, *Class. Quant. Grav.* **29** (2012) 124007, [arXiv:1111.7185].

[321] B. I. et al., *Proposal of the consortium for indian initiative in gravitational-wave observations (indigo)*, *LIGO-India Tech. Rep. No. LIGO-M1100296* (2011).

[322] C. Mills, V. Tiwari, and S. Fairhurst, *Localization of binary neutron star mergers with second and third generation gravitational-wave detectors*, *Phys. Rev.* **D97** (2018), no. 10 104064, [arXiv:1708.00806].

[323] B. P. A. et al., *The lsc-virgo white paper on instrument science (2017-2018 edition)*, *LIGO Technical Report T1700231* (2017).

[324] **LISA** Collaboration, H. Audley et al., *Laser Interferometer Space Antenna*, arXiv:1702.00786.

[325] X. J. Zhu et al., *An all-sky search for continuous gravitational waves in the Parkes Pulsar Timing Array data set*, *Mon. Not. Roy. Astron. Soc.* **444** (2014), no. 4 3709–3720, [arXiv:1408.5129].

[326] C. J. Moore, D. Mihaylov, A. Lasenby, and G. Gilmore, *Astrometric Search Method for Individually Resolvable Gravitational Wave Sources with Gaia*, *Phys. Rev. Lett.* **119** (2017), no. 26 261102, [arXiv:1707.06239].

[327] A. Raccanelli, *Gravitational wave astronomy with radio galaxy surveys*, *Mon. Not. Roy. Astron. Soc.* **469** (2017), no. 1 656–670, [arXiv:1609.09377].

[328] **VIRGO, LIGO Scientific** Collaboration, B. P. Abbott et al., *GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2*, *Phys. Rev. Lett.* **118** (2017), no. 22 221101, [arXiv:1706.01812].

[329] C. M. Will, *Solar system versus gravitational-wave bounds on the graviton mass*, *Class. Quant. Grav.* **35** (2018), no. 17 17LT01, [arXiv:1805.10523].

[330] E. Berti, A. Buonanno, and C. M. Will, *Estimating spinning binary parameters and testing alternative theories of gravity with LISA*, *Phys. Rev.* **D71** (2005) 084025, [gr-qc/0411129].

[331] S. L. Larson and W. A. Hiscock, *Using binary stars to bound the mass of the graviton*, *Phys. Rev.* **D61** (2000) 104008, [gr-qc/9912102].

[332] C. Cutler, W. A. Hiscock, and S. L. Larson, *LISA, binary stars, and the mass of the graviton*, *Phys. Rev.* **D67** (2003) 024015, [gr-qc/0209101].

[333] L. S. Finn and J. D. Romano, *Rømer time-delay determination of the gravitational-wave propagation speed*, *Phys. Rev.* **D88** (2013), no. 2 022001, [arXiv:1304.0369].

[334] S. Mirshekari, N. Yunes, and C. M. Will, *Constraining Generic Lorentz Violation and the Speed of the Graviton with Gravitational Waves*, *Phys. Rev.* **D85** (2012) 024041, [arXiv:1110.2720].

[335] S. I. Vacaru, *Modified Dispersion Relations in Horava-Lifshitz Gravity and Finsler Brane Models*, *Gen. Rel. Grav.* **44** (2012) 1015–1042, [arXiv:1010.5457].

[336] **Virgo, LIGO Scientific** Collaboration, B. P. Abbott et al., *Observation of Gravitational Waves from a Binary Black Hole Merger*, *Phys. Rev. Lett.* **116** (2016), no. 6 061102, [arXiv:1602.03837].

[337] **Virgo, LIGO Scientific** Collaboration, B. P. Abbott et al., *GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence*, *Phys. Rev. Lett.* **116** (2016), no. 24 241103, [arXiv:1606.04855].

[338] N. Yunes, K. Yagi, and F. Pretorius, *Theoretical Physics Implications of the Binary Black-Hole Mergers GW150914 and GW151226*, *Phys. Rev.* **D94** (2016), no. 8 084002, [arXiv:1603.08955].

[339] I. I. Shapiro, *Fourth Test of General Relativity*, *Phys. Rev. Lett.* **13** (1964) 789–791.

[340] L. M. Krauss and S. Tremaine, *Test of the Weak Equivalence Principle for Neutrinos and Photons*, *Phys. Rev. Lett.* **60** (1988) 176.

[341] B. Bertotti, L. Iess, and P. Tortora, *A test of general relativity using radio links with the Cassini spacecraft*, *Nature* **425** (2003) 374–376.

[342] S. Boran, S. Desai, E. O. Kahya, and R. P. Woodard, *GW170817 Falsifies Dark Matter Emulators*, *Phys. Rev.* **D97** (2018), no. 4 041501, [arXiv:1710.06168].

[343] M. A. Green, J. W. Moffat, and V. T. Toth, *Modified Gravity (MOG), the speed of gravitational radiation and the event GW170817/GRB170817A*, *Phys. Lett.* **B780** (2018) 300–302, [arXiv:1710.11177].

[344] T. P. Sotiriou and S.-Y. Zhou, *Black hole hair in generalized scalar-tensor gravity*, *Phys. Rev. Lett.* **112** (2014) 251102, [arXiv:1312.3622].

[345] R. Jackiw and S. Y. Pi, *Chern-Simons modification of general relativity*, *Phys. Rev.* **D68** (2003) 104012, [gr-qc/0308071].

[346] K. Yagi and L. C. Stein, *Black Hole Based Tests of General Relativity*, *Class. Quant. Grav.* **33** (2016), no. 5 054001, [arXiv:1602.02413].

[347] R. Benkel, T. P. Sotiriou, and H. Witek, *Dynamical scalar hair formation around a Schwarzschild black hole*, *Phys. Rev.* **D94** (2016), no. 12 121503, [arXiv:1612.08184].

[348] **Virgo, LIGO Scientific** Collaboration, B. P. Abbott et al., *GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence*, *Phys. Rev. Lett.* **119** (2017), no. 14 141101, [arXiv:1709.09660].

[349] M. Isi and A. J. Weinstein, *Probing gravitational wave polarizations with signals from compact binary coalescences*, arXiv:1710.03794.

[350] M. Isi, M. Pitkin, and A. J. Weinstein, *Probing Dynamical Gravity with the Polarization of Continuous Gravitational Waves*, *Phys. Rev.* **D96** (2017), no. 4 042001, [arXiv:1703.07530].

[351] **VIRGO, LIGO Scientific** Collaboration, J. Aasi et al., *First low frequency all-sky search for continuous gravitational wave signals*, *Phys. Rev.* **D93** (2016), no. 4 042007, [arXiv:1510.03621].

[352] **Virgo, LIGO Scientific** Collaboration, B. P. Abbott et al., *First search for gravitational waves from known pulsars with Advanced LIGO*, *Astrophys. J.* **839** (2017), no. 1 12, [arXiv:1701.07709]. [Erratum: Astrophys. J.851,no.1,71(2017)].

[353] T. Callister, A. S. Biscoveanu, N. Christensen, M. Isi, A. Matas, O. Minazzoli, T. Regimbau, M. Sakellariadou, J. Tasson, and E. Thrane, *Polarization-based Tests of Gravity with the Stochastic Gravitational-Wave Background*, *Phys. Rev.* **X7** (2017), no. 4 041058, [arXiv:1704.08373].

[354] **Virgo, LIGO Scientific** Collaboration, B. P. Abbott et al., *A Search for Tensor, Vector, and Scalar Polarizations in the Stochastic Gravitational-Wave Background*, *Phys. Rev. Lett.* **120** (2018) 201102, [arXiv:1802.10194].

[355] C. M. Will, *Bounding the mass of the graviton using gravitational wave observations of inspiralling compact binaries*, *Phys. Rev.* **D57** (1998) 2061–2068, [gr-qc/9709011].

[356] K. Hinterbichler, M. Trodden, and D. Wesley, *Multi-field galileons and higher co-dimension branes*, *Phys. Rev.* **D82** (2010) 124018, [arXiv:1008.1305].

[357] G. Goon, K. Hinterbichler, and M. Trodden, *Symmetries for Galileons and DBI scalars on curved space*, *JCAP* **1107** (2011) 017, [arXiv:1103.5745].

[358] M. Trodden and K. Hinterbichler, *Generalizing Galileons*, *Class. Quant. Grav.* **28** (2011) 204003, [arXiv:1104.2088].

[359] P. Brax, C. Burrage, and A.-C. Davis, *The Speed of Galileon Gravity*, *JCAP* **1603** (2016), no. 03 004, [arXiv:1510.03701].

[360] J. Beltran Jimenez and A. L. Maroto, *A cosmic vector for dark energy*, *Phys. Rev.* **D78** (2008) 063005, [arXiv:0801.1486].

[361] S. Desai, E. O. Kahya, and R. P. Woodard, *Reduced time delay for gravitational waves with dark matter emulators*, *Phys. Rev.* **D77** (2008) 124041, [arXiv:0804.3804].

[362] R. Kimura, T. Kobayashi, and K. Yamamoto, *Observational Constraints on Kinetic Gravity Braiding from the Integrated Sachs-Wolfe Effect*, *Phys. Rev.* **D85** (2012) 123503, [arXiv:1110.3598].

[363] C. F. B. Macedo, J. Sakstein, E. Berti, L. Gualtieri, H. O. Silva, and T. P. Sotiriou, *Self-interactions and Spontaneous Black Hole Scalarization*, *Phys. Rev.* **D99** (2019), no. 10 104041, [arXiv:1903.06784].

[364] T. Kobayashi, H. Motohashi, and T. Suyama, *Black hole perturbation in the most general scalar-tensor theory with second-order field equations I: the odd-parity sector*, *Phys. Rev.* **D85** (2012) 084025, [arXiv:1202.4893]. [Erratum: Phys. Rev.D96,no.10,109903(2017)].

[365] D. D. Doneva and S. S. Yazadjiev, *New Gauss-Bonnet Black Holes with Curvature-Induced Scalarization in Extended Scalar-Tensor Theories*, *Phys. Rev. Lett.* **120** (2018), no. 13 131103, [arXiv:1711.01187].

[366] H. O. Silva, J. Sakstein, L. Gualtieri, T. P. Sotiriou, and E. Berti, *Spontaneous scalarization of black holes and compact stars from a Gauss-Bonnet coupling*, *Phys. Rev. Lett.* **120** (2018), no. 13 131104, [arXiv:1711.02080].

[367] K. Yagi, *A New constraint on scalar Gauss-Bonnet gravity and a possible explanation for the excess of the orbital decay rate in a low-mass X-ray binary*, *Phys. Rev.* **D86** (2012) 081504, [arXiv:1204.4524].

[368] H. Witek, L. Gualtieri, P. Pani, and T. P. Sotiriou, *Black holes and binary mergers in scalar Gauss-Bonnet gravity: scalar field dynamics*, *Phys. Rev.* **D99** (2019), no. 6 064035, [arXiv:1810.05177].

[369] G. D. Moore and A. E. Nelson, *Lower bound on the propagation speed of gravity from gravitational Cherenkov radiation*, *JHEP* **09** (2001) 023, [hep-ph/0106220].

[370] C. M. Caves, *GRAVITATIONAL RADIATION AND THE ULTIMATE SPEED IN ROSEN'S BIMETRIC THEORY OF GRAVITY*, *Annals Phys.* **125** (1980) 35–52.

[371] R. Kimura and K. Yamamoto, *Constraints on general second-order scalar-tensor models from gravitational Cherenkov radiation*, *JCAP* **1207** (2012) 050, [arXiv:1112.4284].

[372] J. Beltrán Jiménez, F. Piazza, and H. Velten, *Evading the Vainshtein Mechanism with Anomalous Gravitational Wave Speed: Constraints on Modified Gravity from Binary Pulsars*, *Phys. Rev. Lett.* **116** (2016), no. 6 061101, [arXiv:1507.05047].

[373] N. Cornish, D. Blas, and G. Nardini, *Bounding the speed of gravity with gravitational wave observations*, *Phys. Rev. Lett.* **119** (2017), no. 16 161102, [arXiv:1707.06101].

[374] **Virgo, LIGO Scientific** Collaboration, B. Abbott et al., *GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral*, *Phys. Rev. Lett.* **119** (2017), no. 16 161101, [arXiv:1710.05832].

[375] G. Gubitosi and E. V. Linder, *Purely Kinetic Coupled Gravity*, *Phys. Lett.* **B703** (2011) 113–118, [arXiv:1106.2815].

[376] S. Nojiri, S. D. Odintsov, and M. Sasaki, *Gauss-Bonnet dark energy*, *Phys. Rev.* **D71** (2005) 123509, [hep-th/0504052].

[377] T. Baker, E. Bellini, P. G. Ferreira, M. Lagos, J. Noller, and I. Sawicki, *Strong constraints on cosmological gravity from GW170817 and GRB 170817A*, *Phys. Rev. Lett.* **119** (2017), no. 25 251301, [arXiv:1710.06394].

[378] J. Sakstein and B. Jain, *Implications of the Neutron Star Merger GW170817 for Cosmological Scalar-Tensor Theories*, *Phys. Rev. Lett.* **119** (2017), no. 25 251303, [arXiv:1710.05893].

[379] S. Arai and A. Nishizawa, *Generalized framework for testing gravity with gravitational-wave propagation. II. Constraints on Horndeski theory*, *Phys. Rev.* **D97** (2018), no. 10 104038, [arXiv:1711.03776].

[380] L. Amendola, M. Kunz, I. D. Saltas, and I. Sawicki, *Fate of Large-Scale Structure in Modified Gravity After GW170817 and GRB170817A*, *Phys. Rev. Lett.* **120** (2018), no. 13 131101, [arXiv:1711.04825].

[381] A. Emir Gümrükçüoğlu, M. Saravani, and T. P. Sotiriou, *Hořava gravity after GW170817*, *Phys. Rev.* **D97** (2018), no. 2 024032, [arXiv:1711.08845].

[382] Y. Akrami, P. Brax, A.-C. Davis, and V. Vardanyan, *Neutron star merger GW170817 strongly constrains doubly coupled bigravity*, *Phys. Rev.* **D97** (2018), no. 12 124010, [arXiv:1803.09726].

[383] Y.-F. Cai, C. Li, E. N. Saridakis, and L. Xue, *$f(T)$ gravity after GW170817 and GRB170817A*, *Phys. Rev.* **D97** (2018) 103513, [arXiv:1801.05827].

[384] S. Jana, G. K. Chakravarty, and S. Mohanty, *Constraints on Born-Infeld gravity from the speed of gravitational waves after GW170817 and GRB 170817A*, *Phys. Rev.* **D97** (2018), no. 8 084011, [arXiv:1711.04137].

[385] N. Tanahashi and S. Ohashi, *Wave propagation and shock formation in the most general scalar-tensor theories*, *Class. Quant. Grav.* **34** (2017), no. 21 215003, [arXiv:1704.02757].

[386] C. de Rham and A. Matas, *Ostrogradsky in Theories with Multiple Fields*, *JCAP* **1606** (2016), no. 06 041, [arXiv:1604.08638].

[387] N. Franchini and T. P. Sotiriou, *Cosmology with subdominant Horndeski scalar field*, arXiv:1903.05427.

[388] E. J. Copeland, M. Kopp, A. Padilla, P. M. Saffin, and C. Skordis, *Dark energy after GW170817 revisited*, *Phys. Rev. Lett.* **122** (2019), no. 6 061301, [arXiv:1810.08239].

[389] C. de Rham and S. Melville, *Gravitational Rainbows: LIGO and Dark Energy at its Cutoff*, *Phys. Rev. Lett.* **121** (2018), no. 22 221101, [arXiv:1806.09417].

[390] C. Deffayet and K. Menou, *Probing Gravity with Spacetime Sirens*, *Astrophys. J.* **668** (2007) L143–L146, [arXiv:0709.0003].

[391] **Virgo, LIGO Scientific** Collaboration, B. P. Abbott et al., *Tests of General Relativity with GW170817*, arXiv:1811.00364.

[392] M. Lagos, M. Fishbach, P. Landry, and D. E. Holz, *Standard sirens with a running Planck mass*, *Phys. Rev.* **D99** (2019), no. 8 083504, [arXiv:1901.03321].

[393] E. Calabrese, N. Battaglia, and D. N. Spergel, *Testing Gravity with Gravitational Wave Source Counts*, *Class. Quant. Grav.* **33** (2016), no. 16 165004, [arXiv:1602.03883].

[394] K. Pardo, M. Fishbach, D. E. Holz, and D. N. Spergel, *Limits on the number of spacetime dimensions from GW170817*, *JCAP* **1807** (2018), no. 07 048, [arXiv:1801.08160].

[395] E. Belgacem, Y. Dirian, S. Foffa, and M. Maggiore, *Modified gravitational-wave propagation and standard sirens*, *Phys. Rev.* **D98** (2018), no. 2 023510, [arXiv:1805.08731].

[396] R. R. Caldwell and C. Devulder, *Gravitational Wave Opacity from Gauge Field Dark Energy*, arXiv:1802.07371.

[397] D. Comelli, M. Crisostomi, and L. Pilo, *Perturbations in Massive Gravity Cosmology*, *JHEP* **06** (2012) 085, [arXiv:1202.1986].

[398] D. Comelli, M. Crisostomi, F. Nesti, and L. Pilo, *FRW Cosmology in Ghost Free Massive Gravity*, *JHEP* **03** (2012) 067, [arXiv:1111.1983]. [Erratum: JHEP06,020(2012)].

[399] J. Garcia-Bellido, A. D. Linde, and D. Wands, *Density perturbations and black hole formation in hybrid inflation*, *Phys. Rev.* **D54** (1996) 6040–6058, [astro-ph/9605094].

[400] J. Garca-Bellido, *Massive Primordial Black Holes as Dark Matter and their detection with Gravitational Waves*, *J. Phys. Conf. Ser.* **840** (2017), no. 1 012032, [arXiv:1702.08275].

[401] A. Kashlinsky et al., *Electromagnetic probes of primordial black holes as dark matter*, arXiv:1903.04424.

[402] M. Kopp, S. Hofmann, and J. Weller, *Separate Universes Do Not Constrain Primordial Black Hole Formation*, *Phys. Rev.* **D83** (2011) 124025, [arXiv:1012.4369].

[403] T. Harada, C.-M. Yoo, and K. Kohri, *Threshold of primordial black hole formation*, *Phys. Rev.* **D88** (2013), no. 8 084051, [arXiv:1309.4201]. [Erratum: Phys. Rev.D89,no.2,029903(2014)].

[404] S. Young, C. T. Byrnes, and M. Sasaki, *Calculating the mass fraction of primordial black holes*, *JCAP* **1407** (2014) 045, [arXiv:1405.7023].

[405] C.-M. Yoo, T. Harada, J. Garriga, and K. Kohri, *Primordial black hole abundance from random Gaussian curvature perturbations and a local density threshold*, *PTEP* **2018** (2018), no. 12 123E01, [arXiv:1805.03946].

[406] C. Germani and I. Musco, *Abundance of Primordial Black Holes Depends on the Shape of the Inflationary Power Spectrum*, *Phys. Rev. Lett.* **122** (2019), no. 14 141302, [arXiv:1805.04087].

[407] S. Young, I. Musco, and C. T. Byrnes, *Primordial black hole formation and abundance: contribution from the non-linear relation between the density and curvature perturbation*, arXiv:1904.00984.

[408] V. De Luca, G. Franciolini, A. Kehagias, M. Peloso, A. Riotto, and C. nal, *The Ineludible non-Gaussianity of the Primordial Black Hole Abundance*, arXiv:1904.00970.

[409] K. Jedamzik, *Primordial black hole formation during the QCD epoch*, *Phys. Rev.* **D55** (1997) 5871–5875, [astro-ph/9605152].

[410] C. T. Byrnes, M. Hindmarsh, S. Young, and M. R. S. Hawkins, *Primordial black holes with an accurate QCD equation of state*, *JCAP* **1808** (2018), no. 08 041, [arXiv:1801.06138].

[411] J. Garcia-Bellido and E. Ruiz Morales, *Primordial black holes from single field models of inflation*, *Phys. Dark Univ.* **18** (2017) 47–54, [arXiv:1702.03901].

[412] K. Kannike, L. Marzola, M. Raidal, and H. Veermäe, *Single Field Double Inflation and Primordial Black Holes*, *JCAP* **1709** (2017), no. 09 020, [arXiv:1705.06225].

[413] C. Germani and T. Prokopec, *On primordial black holes from an inflection point*, *Phys. Dark Univ.* **18** (2017) 6–10, [arXiv:1706.04226].

[414] H. Motohashi and W. Hu, *Primordial Black Holes and Slow-Roll Violation*, *Phys. Rev.* **D96** (2017), no. 6 063503, [arXiv:1706.06784].

[415] G. Ballesteros and M. Taoso, *Primordial black hole dark matter from single field inflation*, *Phys. Rev.* **D97** (2018), no. 2 023501, [arXiv:1709.05565].

[416] M. Cicoli, V. A. Diaz, and F. G. Pedro, *Primordial Black Holes from String Inflation*, *JCAP* **1806** (2018), no. 06 034, [arXiv:1803.02837].

[417] O. zsoy, S. Parameswaran, G. Tasinato, and I. Zavala, *Mechanisms for Primordial Black Hole Production in String Theory*, *JCAP* **1807** (2018) 005, [arXiv:1803.07626].

[418] I. Dalianis, A. Kehagias, and G. Tringas, *Primordial black holes from -attractors*, *JCAP* **1901** (2019) 037, [arXiv:1805.09483].

[419] S. M. Leach and A. R. Liddle, *Inflationary perturbations near horizon crossing*, *Phys. Rev.* **D63** (2001) 043508, [astro-ph/0010082].

[420] S. M. Leach, M. Sasaki, D. Wands, and A. R. Liddle, *Enhancement of superhorizon scale inflationary curvature perturbations*, *Phys. Rev.* **D64** (2001) 023512, [astro-ph/0101406].

[421] V. F. Mukhanov, *Gravitational Instability of the Universe Filled with a Scalar Field*, *JETP Lett.* **41** (1985) 493–496. [Pisma Zh. Eksp. Teor. Fiz.41,402(1985)].

[422] M. Sasaki, *Large Scale Quantum Fluctuations in the Inflationary Universe*, *Prog. Theor. Phys.* **76** (1986) 1036.

[423] G. Ballesteros, J. Beltran Jimenez, and M. Pieroni, *Black hole formation from a general quadratic action for inflationary primordial fluctuations*, *JCAP* **1906** (2019), no. 06 016, [arXiv:1811.03065].

[424] R. Saito, J. Yokoyama, and R. Nagata, *Single-field inflation, anomalous enhancement of superhorizon fluctuations, and non-Gaussianity in primordial black hole formation*, *JCAP* **0806** (2008) 024, [arXiv:0804.3470].

[425] R. Allahverdi, K. Enqvist, J. Garcia-Bellido, and A. Mazumdar, *Gauge invariant MSSM inflaton*, *Phys. Rev. Lett.* **97** (2006) 191304, [hep-ph/0605035].

[426] M. H. Namjoo, H. Firouzjahi, and M. Sasaki, *Violation of non-Gaussianity consistency relation in a single field inflationary model*, *EPL* **101** (2013), no. 3 39001, [arXiv:1210.3692].

[427] J. Martin, H. Motohashi, and T. Suyama, *Ultra Slow-Roll Inflation and the non-Gaussianity Consistency Relation*, *Phys. Rev.* **D87** (2013), no. 2 023514, [arXiv:1211.0083].

[428] X. Chen, H. Firouzjahi, M. H. Namjoo, and M. Sasaki, *A Single Field Inflation Model with Large Local Non-Gaussianity*, *EPL* **102** (2013), no. 5 59001, [arXiv:1301.5699].

[429] C. Pattison, V. Vennin, H. Assadullahi, and D. Wands, *The attractive behaviour of ultra-slow-roll inflation*, *JCAP* **1808** (2018), no. 08 048, [arXiv:1806.09553].

[430] **Planck** Collaboration, P. A. R. Ade et al., *Planck 2015 results. XX. Constraints on inflation*, *Astron. Astrophys.* **594** (2016) A20, [arXiv:1502.02114].

[431] F. L. Bezrukov and M. Shaposhnikov, *The Standard Model Higgs boson as the inflaton*, *Phys. Lett.* **B659** (2008) 703–706, [arXiv:0710.3755].

[432] K. Allison, *Higgs xi-inflation for the 125-126 GeV Higgs: a two-loop analysis*, *JHEP* **02** (2014) 040, [arXiv:1306.6931].

[433] **Particle Data Group** Collaboration, M. Tanabashi et al., *Review of Particle Physics*, *Phys. Rev.* **D98** (2018), no. 3 030001.

[434] **ATLAS** Collaboration, M. Aaboud et al., *Measurement of the Higgs boson mass in the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector*, *Phys. Lett.* **B784** (2018) 345–366, [arXiv:1806.00242].

[435] **CMS** Collaboration, C. Collaboration, *Measurement of the top quark mass with lepton+jets final states in pp collisions at $\sqrt{s} = 13$ TeV*, .

[436] F. Bezrukov, M. Yu. Kalmykov, B. A. Kniehl, and M. Shaposhnikov, *Higgs Boson Mass and New Physics*, *JHEP* **10** (2012) 140, [arXiv:1205.2893]. [,275(2012)].

[437] F. Bezrukov and M. Shaposhnikov, *Higgs inflation at the critical point*, *Phys. Lett.* **B734** (2014) 249–254, [arXiv:1403.6078].

[438] Y. Hamada, H. Kawai, K.-y. Oda, and S. C. Park, *Higgs Inflation is Still Alive after the Results from BICEP2*, *Phys. Rev. Lett.* **112** (2014), no. 24 241301, [arXiv:1403.5043].

[439] M. Herranen, T. Markkanen, S. Nurmi, and A. Rajantie, *Spacetime curvature and the Higgs stability during inflation*, *Phys. Rev. Lett.* **113** (2014), no. 21 211102, [arXiv:1407.3141].

[440] F. Bezrukov, J. Rubio, and M. Shaposhnikov, *Living beyond the edge: Higgs inflation and vacuum metastability*, *Phys. Rev.* **D92** (2015), no. 8 083512, [arXiv:1412.3811].

[441] V.-M. Enckell, K. Enqvist, and S. Nurmi, *Observational signatures of Higgs inflation*, *JCAP* **1607** (2016), no. 07 047, [arXiv:1603.07572].

[442] F. Bezrukov, M. Pauly, and J. Rubio, *On the robustness of the primordial power spectrum in renormalized Higgs inflation*, *JCAP* **1802** (2018), no. 02 040, [arXiv:1706.05007].

[443] A. Salvio, *Critical Higgs inflation in a Viable Motivated Model*, *Phys. Rev.* **D99** (2019), no. 1 015037, [arXiv:1810.00792].

[444] **Planck** Collaboration, Y. Akrami et al., *Planck 2018 results. X. Constraints on inflation*, arXiv:1807.06211.

[445] I. Masina, *Ruling out Critical Higgs Inflation?*, *Phys. Rev.* **D98** (2018), no. 4 043536, [arXiv:1805.02160].

[446] J. Garcia-Bellido and D. G. Figueroa, *A stochastic background of gravitational waves from hybrid preheating*, *Phys. Rev. Lett.* **98** (2007) 061302, [astro-ph/0701014].

[447] J. R. Chisholm, *Clustering of primordial black holes: basic results*, *Phys. Rev.* **D73** (2006) 083504, [astro-ph/0509141].

[448] Y. Ali-Hamoud, *Correlation Function of High-Threshold Regions and Application to the Initial Small-Scale Clustering of Primordial Black Holes*, *Phys. Rev. Lett.* **121** (2018), no. 8 081304, [arXiv:1805.05912].

[449] V. Desjacques and A. Riotto, *Spatial clustering of primordial black holes*, *Phys. Rev.* **D98** (2018), no. 12 123533, [arXiv:1806.10414].

[450] J. Garca-Bellido and S. Clesse, *Constraints from microlensing experiments on clustered primordial black holes*, *Phys. Dark Univ.* **19** (2018) 144–148, [arXiv:1710.04694].

[451] S. Chandrasekhar, *The maximum mass of ideal white dwarfs*, *Astrophys. J.* **74** (1931) 81–82.

[452] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott et al., *Search for Subsolar-Mass Ultracompact Binaries in Advanced LIGOs First Observing Run*, *Phys. Rev. Lett.* **121** (2018), no. 23 231103, [arXiv:1808.04771].

[453] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott et al., *Search for sub-solar mass ultracompact binaries in Advanced LIGO's second observing run*, arXiv:1904.08976.

[454] N. Fernandez and S. Profumo, *Unraveling the origin of black holes from effective spin measurements with LIGO-Virgo*, arXiv:1905.13019.

[455] S. Clesse and J. Garca-Bellido, *Detecting the gravitational wave background from primordial black hole dark matter*, *Phys. Dark Univ.* **18** (2017) 105–114, [arXiv:1610.08479].

[456] N. Bartolo et al., *Science with the space-based interferometer LISA. IV: Probing inflation with gravitational waves*, *JCAP* **1612** (2016), no. 12 026, [arXiv:1610.06481].

[457] J. Garcia-Bellido, M. Peloso, and C. Unal, *Gravitational Wave signatures of inflationary models from Primordial Black Hole Dark Matter*, *JCAP* **1709** (2017), no. 09 013, [arXiv:1707.02441].

[458] M. Drees and Y. Xu, *Critical Higgs Inflation and Second Order Gravitational Wave Signatures*, arXiv:1905.13581.

[459] B. Carr, M. Raidal, T. Tenkanen, V. Vaskonen, and H. Veermae, *Primordial black hole constraints for extended mass functions*, *Phys. Rev.* **D96** (2017), no. 2 023514, [arXiv:1705.05567].

[460] J. Garcia-Bellido, S. Clesse, and P. Fleury, *Primordial black holes survive SN lensing constraints*, *Phys. Dark Univ.* **20** (2018) 95–100, [arXiv:1712.06574].

[461] A. A. Starobinsky, *STOCHASTIC DE SITTER (INFLATIONARY) STAGE IN THE EARLY UNIVERSE*, *Lect. Notes Phys.* **246** (1986) 107–126.

[462] A. J. Tolley and M. Wyman, *Stochastic Inflation Revisited: Non-Slow Roll Statistics and DBI Inflation*, *JCAP* **0804** (2008) 028, [arXiv:0801.1854].

[463] J. Grain and V. Vennin, *Stochastic inflation in phase space: Is slow roll a stochastic attractor?*, *JCAP* **1705** (2017), no. 05 045, [arXiv:1703.00447].

[464] S. Matarrese, F. Lucchin, and S. A. Bonometto, *A Path Integral Approach To Large Scale Matter Distribution Originated By Nongaussian Fluctuations*, *Astrophys. J.* **310** (1986) L21–L26.

[465] C. T. Byrnes, E. J. Copeland, and A. M. Green, *Primordial black holes as a tool for constraining non-Gaussianity*, *Phys. Rev.* **D86** (2012) 043512, [arXiv:1206.4188].

[466] S. Young and C. T. Byrnes, *Primordial black holes in non-Gaussian regimes*, *JCAP* **1308** (2013) 052, [arXiv:1307.4995].

[467] S. Young, D. Regan, and C. T. Byrnes, *Influence of large local and non-local bispectra on primordial black hole abundance*, *JCAP* **1602** (2016), no. 02 029, [arXiv:1512.07224].

[468] Y. Tada and S. Yokoyama, *Primordial black holes as biased tracers*, *Phys. Rev.* **D91** (2015), no. 12 123534, [arXiv:1502.01124].

[469] S. Young and C. T. Byrnes, *Signatures of non-gaussianity in the isocurvature modes of primordial black hole dark matter*, *JCAP* **1504** (2015), no. 04 034, [arXiv:1503.01505].

[470] G. Franciolini, A. Kehagias, S. Matarrese, and A. Riotto, *Primordial Black Holes from Inflation and non-Gaussianity*, *JCAP* **1803** (2018), no. 03 016, [arXiv:1801.09415].

[471] J. S. Bullock and J. R. Primack, *NonGaussian fluctuations and primordial black holes from inflation*, *Phys. Rev.* **D55** (1997) 7423–7439, [astro-ph/9611106].

[472] P. Ivanov, *Nonlinear metric perturbations and production of primordial black holes*, *Phys. Rev.* **D57** (1998) 7145–7154, [astro-ph/9708224].

[473] J. Yokoyama, *Chaotic new inflation and formation of primordial black holes*, *Phys. Rev.* **D58** (1998) 083510, [astro-ph/9802357].

[474] C. Pattison, V. Vennin, H. Assadullahi, and D. Wands, *Quantum diffusion during inflation and primordial black holes*, *JCAP* **1710** (2017), no. 10 046, [arXiv:1707.00537].

[475] M. Biagetti, G. Franciolini, A. Kehagias, and A. Riotto, *Primordial Black Holes from Inflation and Quantum Diffusion*, *JCAP* **1807** (2018), no. 07 032, [arXiv:1804.07124].

[476] A. D. Linde, D. A. Linde, and A. Mezhlumian, *From the Big Bang theory to the theory of a stationary universe*, *Phys. Rev.* **D49** (1994) 1783–1826, [gr-qc/9306035].

[477] C. Pattison, V. Vennin, H. Assadullahi, and D. Wands, *Stochastic inflation beyond slow roll*, arXiv:1905.06300.

[478] V. Vennin and A. A. Starobinsky, *Correlation Functions in Stochastic Inflation*, *Eur. Phys. J.* **C75** (2015) 413, [arXiv:1506.04732].

[479] S. Winitzki and A. Vilenkin, *Effective noise in stochastic description of inflation*, *Phys. Rev.* **D61** (2000) 084008, [gr-qc/9911029].

[480] H. Assadullahi, H. Firouzjahi, M. Noorbala, V. Vennin, and D. Wands, *Multiple Fields in Stochastic Inflation*, *JCAP* **1606** (2016), no. 06 043, [arXiv:1604.04502].

[481] J. Garcia-Bellido, M. Peloso, and C. Unal, *Gravitational waves at interferometer scales and primordial black holes in axion inflation*, *JCAP* **1612** (2016), no. 12 031, [arXiv:1610.03763].

[482] L. Lombriser and A. Taylor, *Breaking a Dark Degeneracy with Gravitational Waves*, *JCAP* **1603** (2016), no. 03 031, [arXiv:1509.08458].

[483] M. Ishak et al., *Modified Gravity and Dark Energy models Beyond $w(z)$CDM Testable by LSST*, arXiv:1905.09687.

[484] K. Bechtol et al., *Dark Matter Science in the Era of LSST*, arXiv:1903.04425.

[485] L. Amendola et al., *Cosmology and fundamental physics with the Euclid satellite*, *Living Rev. Rel.* **21** (2018), no. 1 2, [arXiv:1606.00180].

[486] O. Dor et al., *WFIRST: The Essential Cosmology Space Observatory for the Coming Decade*, arXiv:1904.01174.

[487] P. Creminelli, M. Lewandowski, G. Tambalo, and F. Vernizzi, *Gravitational Wave Decay into Dark Energy*, *JCAP* **1812** (2018), no. 12 025, [arXiv:1809.03484].

[488] V. Kalogera et al., *Deeper, Wider, Sharper: Next-Generation Ground-Based Gravitational-Wave Observations of Binary Black Holes*, arXiv:1903.09220.

[489] A. Pirkowska, M. Biesiada, and Z.-H. Zhu, *Strong gravitational lensing of gravitational waves in Einstein Telescope*, *JCAP* **1310** (2013) 022, [arXiv:1309.5731].

[490] A. Sesana, *The promise of multi-band gravitational wave astronomy*, *Phys. Rev. Lett.* **116** (2016), no. 23 231102, [arXiv:1602.06951].

[491] A. De Felice, L. Heisenberg, R. Kase, S. Mukohyama, S. Tsujikawa, and Y.-l. Zhang, *Cosmology in generalized Proca theories*, *JCAP* **1606** (2016), no. 06 048, [arXiv:1603.05806].

[492] P. M. Chesler and A. Loeb, *Constraining Relativistic Generalizations of Modified Newtonian Dynamics with Gravitational Waves*, *Phys. Rev. Lett.* **119** (2017), no. 3 031102, [arXiv:1704.05116].

[493] T. Ortín, *Gravity and strings*. Cambridge University Press, 2004.

[494] K. Van Acoleyen and J. Van Doorsselaere, *Galileons from Lovelock actions*, *Phys. Rev.* **D83** (2011) 084025, [arXiv:1102.0487].

[495] J. R. Gair, M. Vallisneri, S. L. Larson, and J. G. Baker, *Testing general relativity with low-frequency, space-based gravitational-wave detectors*, *Living Reviews in Relativity* **16** (2013), no. 7.

[496] A. Stroeer and A. Vecchio, *The LISA verification binaries*, *Class. Quant. Grav.* **23** (2006) S809–S818, [astro-ph/0605227].

[497] P. Amaro-Seoane et al., *Low-frequency gravitational-wave science with eLISA/NGO*, *Class. Quant. Grav.* **29** (2012) 124016, [arXiv:1202.0839].

[498] G. Nelemans, "Lisa verification binaries."

[499] W. R. Brown, M. Kilic, J. J. Hermes, C. Allende Prieto, S. J. Kenyon, and D. E. Winget, *A 12 minute Orbital Period Detached White Dwarf Eclipsing Binary*, *Astrophys. J.* **737** (2011) L23, [arXiv:1107.2389].

[500] J. J. Hermes, M. Kilic, W. R. Brown, D. E. Winget, C. Allende Prieto, A. Gianninas, A. S. Mukadam, A. Cabrera-Lavers, and S. J. Kenyon, *Rapid Orbital Decay in the 12.75-minute WD+WD Binary J0651+2844*, *Astrophys. J.* **757** (2012) L21, [arXiv:1208.5051].

[501] M. Vallisneri, *Use and abuse of the Fisher information matrix in the assessment of gravitational-wave parameter-estimation prospects*, *Phys. Rev.* **D77** (2008) 042001, [gr-qc/0703086].